Promoting Healthy Families
Data Management Plan

Created by: Amber Rieder, Natalie Harvey-Younis, Yash Joshi, & Venessa Yeung

STRoNG Families

The Tri-Agency policy states that:

"...all DMPs should describe:

- how data will be collected, documented, formatted, protected and preserved;
- how existing datasets will be used and what new data will be created over the course of the research project;
- whether and how data will be shared; and
- where data will be deposited.

DMPs also indicate who is responsible for managing the project's data, describe the succession plans in place should that person leave the research team, and identify the data-related roles and responsibilities of other team members where appropriate. Finally, DMPs outline ethical, legal and commercial constraints the data are subject to, and methodological considerations that support or preclude data sharing."

Table of Contents

Administrative Details

Project Title: Promoting Health Families

Institution: McMaster University; Offord Center for Child Studies

Principal Investigator (PI): Dr. Andrea Gonzalez

Funder: Public Health Agency of Canada

Project Abstract: Promoting Healthy Families: A Canadian Evaluation of Two Evidence-based Parenting Programs – aims to rigorously implement and evaluate the effectiveness of Triple P – Positive Parenting Program (Triple P) and Circle of Security Parenting Program (COSP), in promoting healthy family relationships, positive child outcomes and preventing child maltreatment. Specifically, the project will assess whether Triple P and COSP respectively will reduce: 1) child maltreatment-related outcomes; 2) improve parenting practices and functioning and; 3) improve child developmental outcomes.

Start of funding: April 1, 2018

End of funding: March 30, 2024 (with the recent extension/amendment)

Last modified: April 26, 2023

Plan Overview

The Promoting Healthy Families (PHF) Data Management Plan (DMP) is a living research document. Team members have access to the document and should update it frequently. The plan may be used as a check for what's going well, and what needs changing, as well as articulate team roles and security measures as it pertains to data management.

During the initial drafting of the DMP, the DMT conducted yearly review meetings with McMaster Data Management Services Team member, Danica Evering (everingd@mcmaster.ca). Consultations regarding the DMP will continue on an as-needed basis.

Team Roles

Table 1: Team roles, contact information, and start-end dates

Name	Role	Contact	Start-End Date	ORCID ID

List of Acronyms

ACE | Adverse Childhood Experiences
AE | Adverse Event
ASP | Application Service Provider
AUDIT | Alcohol Use Disorders Identification Test
AUDIT-C | Alcohol USe Disorders Identification Test Consumption
CA-SUS | Child and Adolescent Use Schedule
CCQ | Composite Caregiving Questionnaire
CES-D | Center for Epidemiologic Studies Depression Scale (20-item)
CES-D-10 | Centre for Epidemiological Depression Scale (10-item)
CHEERS | Consolidated Health Economic Evaluation Reporting Standards
CIHI | Canadian Institute for Health Information
CSV | Comma Separated Values
CIS | Canadian Incidence Study of Reported Child Abuse and Neglect
COSP | Circle of Security Parenting Program
CTQ | Childhood Trauma Questionnaire
CTSPC-R | Child Version, Parent-Child Conflict Tactics Scale - Revised
DAD | Discharge Abstract Database
DAS-7 | Abbreviated Dyadic Adjustment Scale
DAS-32 | Dyadic Adjustment Scale
DERS-SF | Difficulties in Emotion Regulation Scale - Short Form
DMP | Data Management Plan
DMT |Data Management Team
DSA | Data Sharing Agreement
DSMB | Data Safety Monitoring Board
EAS | Emotional Availability Scale
ECG | Electrocardiograph
EDA | Electrodermal Skin Conductance
GAD | Generalized Anxiety Disorder
GAD-7 | Generalized Anxiety Disorder Scale
HRQOL | Health-Related Quality of Life
HRV | Heart Rate Variability
HSCS-PS | Health Status Classification System Preschool Version
HUI3 | Health Unity Index - Mark 3
IBI | Inter-Beat Intervals
ICD | International Classification of Diseases
ICES | Integrative Clinical Evaluative Sciences
ID | Identification
IPV | Intimate Partner Violence

LS | Laxness Scale
LX | Laxness
NACRS | National Ambulatory Care Reporting System
NICU | Neonatal Intensive Care Unit
OHIP | Ontario Health Insurance Plan
OR | Overreactivity
OS | Overreactivity Scale
PCV | Plotkin Child Vignettes
PHF | Promoting Healthy Families
PHIPA | Personal Health Information Protection Act
PI | Principal Investigator
PIPEDA | Personal Information Protection and Electronic Documents Act
PS | Parenting Scale
QALYs | Quality Adjusted Life Years
RA | Research Assistant
RC | Research Coordinator
RCT | Randomised Control Trial
REB | Research Ethics Board
SaaS | Software-as-a-Service
SCL | Skin Conductance Level
SAS | Statistical Analysis System
SDQ | Strengths and Difficulties Questionnaire
SOP | Standard Operating Procedure
SPSS | Statistical Package for Social Sciences
SSL VPN | Secure Sockets Layer Virtual Private Network
TAU | Treatment as Usual
TLS | Transport Layer Security
VPN | Virtual Private Network
VB | Verbosity
VS | Verbosity Scale
WAI | Working Alliance Inventory
WAI-SR | Working Alliance Inventory-Short Revised
WHO | World Health Organization

Child Maltreatment and Parenting Programs in Canada

- Child maltreatment is a significant global public health problem with an estimated incidence of over 25% of children experiencing one or more forms of maltreatment in their childhood.
- The Canadian Incidence Study of Reported Child Abuse and Neglect (CIS; 2008) found an incidence of ~39 per 1,000 children for maltreatment-related investigations, and ~14 per 1,000 children for substantiated maltreatment.

The primary categories of substantiated maltreatment were physical abuse (20%), neglect (34%), sexual abuse (3%), emotional abuse (9%), and exposure to intimate partner violence (IPV; 34%).

- Official statistics seriously underestimate the occurrence of child maltreatment, with the annual financial burden estimated in the billions of dollars and maltreatment being a leading cause of health inequality and social injustice.
- Child maltreatment is related to parenting and risk factors affecting children and families, including parenting stress, partner conflicts, lack of social support, caregiver mental health problems, and parental history of abuse/foster care.

Parenting programs have the potential to improve emotional and behavioural adjustment of children, enhance the psychosocial well-being of parents, reduce risk factors and increase protective factors associated with maltreatment.

- However, most parenting programs currently implemented in Canada are inadequately evaluated or not evaluated at all. The authors plan to rigorously evaluate the effectiveness of two parenting programs, Triple P and Circle of Security Parenting Program, in promoting healthy family relationships and preventing child maltreatment. Both show promise, but require further evaluation.

Introduction

Child maltreatment represents a significant public health problem with global meta-analyses suggesting that more than a quarter of children have experienced one or more of these types of maltreatment in their childhood (Stoltenborgh et al., 2012; Stoltenborgh et al., 2013). In Canada, the Canadian Incidence Study of Reported Child Abuse and Neglect has provided information about the estimated number of maltreatment-related investigations over three waves of data collection. Findings from the most recent cycle (2008), indicate that the incidence was 39.16 per 1,000 children; the incidence for substantiated maltreatment was 14.19 per 1,000 children (Public Health Agency of Canada, 2010). The primary categories of substantiated maltreatment based on CIS-2008 data were as follows: physical abuse (20%), sexual abuse (3%), neglect (34%), emotional abuse (9%), and exposure to IPV (34%).

It is widely recognized that these official statistics seriously underestimate the occurrence of child maltreatment. Recent Canadian cost estimates indicate the annual financial burden of family violence in the billions of dollars (Bowlus et al., 2003). In addition to this economic impact, child maltreatment is a leading cause of health inequality and social injustice (WHO, 2006). The potential health and social costs of maltreatment are considerable, with maltreatment being associated with a broad range of negative outcomes including injuries and fatalities, social-emotional and behavioural problems, academic, cognitive and interpersonal difficulties, and increased risk for mental and physical health outcomes (Afifi et al., 2014; Gonzalez et al., 2012; Maguire et al., 2015; Norman et al., 2012; PHAC, 2016; Romano et al., 2015). Meta-analyses and systematic reviews also suggest that exposure to child maltreatment may increase the risk of gender-based violence across the lifespan, including future IPV perpetration or victimization (Kimber et al., 2018; Park & Kim, 2018).

Child maltreatment is also integrally related to parenting and to many of the risk factors that affect children and families including, parenting stress, partner conflicts, lack of social support, caregiver mental health problems, and parental history of abuse/foster care (Gonzalez & MacMillan, 2008; PHAC, 2010). Anonymous surveys from parents suggest that parenting practices that could conceivably be considered abusive (e.g. harsh discipline) are many times more prevalent than official records indicate (Theodore et al., 2005). Thus, given the high rates of self-reports of child maltreatment, data suggesting that perpetrators are typically parents or other adult caregivers (Finkelhor et al., 2014), and that many cases of harsh parenting are never officially reported, an examination of parenting interventions that target parenting challenges and specific risk factors associated with child maltreatment, and can be broadly applied, is greatly needed. Indeed, systematic reviews and meta-analyses suggest that parenting programs can improve emotional and behavioural adjustment of

children, enhance the psychosocial well-being of parents, and may reduce risk factors and increase protective factors associated with maltreatment (Barlow & Coren, 2018; Barlow et al., 2002; Chen & Chan, 2016). A recent posting of the draft recommendation statement from the US Preventive Services Task Force made available for public comment concludes that "the current evidence is insufficient to assess the balance of benefits and harms of primary care interventions to prevent child maltreatment" (URL). These well-documented consequences of child maltreatment, and potential benefits of parenting programs, yet lack of existing evidence about their ability to prevent child maltreatment, emphasize the urgent need to identify and evaluate evidence-based interventions for the prevention of child maltreatment with the potential for nationwide scale-up in Canada (PHAC, 2016).

In 2009, the World Health Organization (WHO) identified interventions that promote safe, stable and nurturing relationships between caregivers and children as key to reducing child maltreatment (WHO, 2009). Although a broad range of parenting programs are currently implemented in communities across Canada, the overwhelming majority are inadequately evaluated, or else not evaluated at all. Furthermore, several parent-training programs are used with the stated goal of preventing child maltreatment; however, few have undergone rigorous evaluations using official maltreatment indicators as outcomes. Too often, interventions are implemented before undergoing adequate evaluation and the term 'promising' is interpreted as sufficient for widespread dissemination (MacMillan et al., 2009). This knowledge gap is an important problem. As noted in the It's Time: Canada's Strategy to Prevent and Address Gender-based Violence, implementing and evaluating programs to prevent gender-based violence, including child maltreatment is a priority for the Government of Canada.

For these reasons, we plan to address the important evidence gaps and rigorously evaluate the effectiveness of two parenting programs, the Triple P – Positive Parenting Program (hereafter referred to as Triple P) and the Circle of Security Parenting Program (hereafter referred to as COSP), in promoting healthy family relationships and preventing child maltreatment. To date, one population-level parenting program, the Triple P, has shown promise in preventing child abuse and neglect, however, it requires further assessment and replication (Altafim & Linhares, 2016; MacMillan et al., 2009; PHAC, 2016; Poole et al., 2014; Prinz et al., 2009; 2016; Sanders et al., 2008; Zemp et al., 2016). Another program, the COSP, aimed at improving insecure and disorganised attachments in children (COS; Cooper et al., 2005; Marvin et al., 2002) has a growing body of literature (Yaholkoski et al., 2016) and while evidence for its' efficacy is promising, further studies are needed. To date there have not been any studies that have focused specifically on the prevention of child abuse and neglect.

Data Collection & Safety Measures

What types of data will be collected or created?

a. Screening data (questionnaires) for randomised control trial (RCT) inclusion/exclusion
b. Personal data pertaining to informed consent
c. Mailing address to send study materials (e.g. workbooks, iPads)
d. Email address for study-related communication (e.g. zoom meeting links, scheduling)
e. Random assignment spreadsheets; agency/group/facilitator/caregiver
f. Questionnaires (demographic data, caregiver and child wellbeing, dyadic/interpersonal relationship quality, parenting (style, attitudes, behaviour, mental and physical health, stress, childhood adversity, emotion regulation, therapeutic alliance, fidelity ratings)
g. Computer-based tasks & output
h. Video-taped parent-child interactions & coding
i. Psychophysiological measures
    i. Electrocardiograph (ECG)
    ii. Electrodermal skin conductance (EDA)
j. Intervention audio recordings
k. Intervention attendance/dosage
l. Interviews and Focus Groups transcripts
m. Data analysis files (e.g. spreadsheets, qualitative memos)
n. Integrative Clinical Evaluative Sciences (ICES) data linkage data

See estimated sample sizes in Table 2 below.

Table 2: Estimated Sample Sizes

Participant	Data	Sample size
Caregivers (screened)	*Quantitative:* Questionnaires	N=
Caregivers (randomized)	*Quantitative:* Questionnaires; Computer-based tasks; Video-taped interactions; group attendance	n=
Caregivers (subset)	*Qualitative:* Interviews and focus groups	n= ●
Caregivers (subset)	*Psychophysiological measures:* ECG, EDA	n=
Providers (agency)	*Quantitative:* Questionnaires (e.g. Readiness)	
Providers (subset)	*Qualitative:* Interviews and focus groups	n= ●
Providers (group facilitators)	*Implementation:* Group & Session Fidelity	n=
Other	Triple P Group Session Fidelity Audio Evaluations	n=1

Figure 1: CONSORT Diagram

How will the data be collected or created?

Study Visit Procedures:

Trained research assistants (RAs) will gather data from all participants during four study assessment time points. Prior to COVID-19, visits would have occurred during in-person, home visits; however, we have modified the protocol and the baseline and post-treatment visits will occur virtually and the 6- and 12-month follow-up visits will occur in-person. If physical distancing measures are still in place at this later date, we will prepare another ethics amendment to reflect remote, virtual data collection procedures. The date and time of visits will be scheduled at the participant's convenience.

Questionnaires and computer-based tasks will all be available online using a web-based link (Qualtrics) and videotaped parent-child interactions will take place over Zoom for baseline and post-treatment, and in-person for follow-up visits. If the 12-month follow-up visit occurs in-person, psychophysiological measures (electrocardiograph (ECG) and electrodermal (EDA; measured as skin conductance) will be collected during the parent-child interactions from a subsample of participants. Once the study is underway, participants will be reminded of study visits via their preferred mode of communication. Assessments will take approximately 60-90 minutes depending on the visit time point. Caregivers will be contacted by phone at 3- and 9-month follow-up time points for a brief 'check-in' session regarding verification of contact information and questions regarding healthcare use and use of other community services in the past three months. These calls will not take more than 15 minutes to complete.

The types of data collected and the file storage locations are listed below in Table 2. Data storage is managed in accordance with McMaster University's Document Storage Guidelines.

Table 3: Data types, sources, formats and size

Data Types	Data Sources	File Formats	Software	Size
Participant Tracking (Visits dates/completions, Consent Files, Contact Information)	*Computers*	*.csv,.xlsx,.xml, .pdf, .docx, .sps, .sav, .spv, .r, .rdata, .rds, .html*	*Qualtrics, SPSS, R, Excel.*	*3.13MB*
Questionnaires (demographic data, caregiver and child wellbeing, dyadic/interpersonal relationship quality, parenting (style, attitudes, behaviour, mental and physical health, stress, childhood adversity, emotion regulation, therapeutic alliance, fidelity ratings)	*Mobile devices (phones, tablets, computers)*	*.csv,.xlsx,.xml, .pdf, .docx, .sps, .sav, .spv, .r, .rdata, .rds, .html*	*Qualtrics, SPSS, R, Excel.*	*ex. total ~5GB*
Audio Recordings (intervention group recordings, interviews, focus groups) and transcripts	*Mobile devices (phones, tablets, computers)*	*.mp3,.mp4, .m4a,.txt, .vtt, .m3u, .docx, .xlsx*	*Zoom, Quicktime, NVivo*	*50.88GB for group audio recordings*
Random assignment and participant key (agency, group), contact information, and attendance tracking	*Mobile devices (phones, tablets, computers)*	*.csv,.xlsx,.docx*	*Excel, Word*	*16.73MB*
File Transfer Protocol Software (secured software to transfer Zoom audio files to transcription service)	*Mobile devices (phones, tablets, computers)*	*.m4a (encrypted)*	*Sync.com*	*Total 1.0 TB*
Computer-based tasks	*Mobile devices (phones, tablets, computers)*	*.csv,.xlsx,.xml, .pdf, .docx, .sps, .sav, .spv, .r, .rdata, .rds, .html*	*Qualtrics, SPSS, R, Excel.*	
Video-taped parent-child interactions & coding	*Mobile devices (phones, tablets, computers)* *Video-recording devices*	*.csv,.xlsx,.xml, .pdf, .docx, .sps, .sav, .spv, .r, .rdata, .rds, .html, . mov., mp4,. mp3*	*Qualtrics, SPSS, R, Excel.*	*306.36GB*
Psychophysiological measures, including time-series data of inter-beat interval (IBI), respiratory sinus arrhythmia (RSA), and skin conductance level (SCL)	*Electrocardiograph (ECG) and Electrodermal (EDA) signals collected during interaction tasks*	*.acq (raw data), .xlsx (cleaned output data), .csv (processed data)*	*Acqknowledge, Mindware ,R*	*Total 898MB*

How will data be accessed and stored?

Data Access:

The Research Coordinator (RC), in collaboration with the Principal Investigators (PIs), will oversee all aspects of the data collection and management at the Offord Centre for Child Studies at McMaster Innovation Park. For the purposes of this study, data is separated into three categories: participant treatment allocation; identifying information (see section below); and de-identified research data (unique participant code only). Access to each of these categories of data will be limited to certain research team members in the following ways:

Participant treatment allocation: The list matching all participants' unique codes with their treatment allocation will be available only to the PIs, graduate and postdoctoral trainees, and the RC. RAs/field interviewers will not have access to participant treatment allocation.

Identifying information: Only the PIs, RC, and DMT will have access to the full list of direct identifiers (including Ontario Health Insurance Plan (OHIP) numbers). The 'key' that matches all participants' unique study codes back to the full list of direct identifiers will be stored in secure and encrypted electronic formats within locked office of the lead PI, Andrea Gonzalez (if working on site), or fully encrypted and password protected computer of the lead PI, Andrea Gonzalez (working offsite). RAs and providers will continue to have access to some identifying information (including name, telephone number, email address, and address) for interview and program purposes. Each organisation will maintain their screening list and client files, detailing the names and telephone numbers of all participants screened for eligibility.

De-identified research data: A senior Study Team member (e.g., Data Manager) designated by the PIs and RC will oversee storage, quality control, and management of de-identified research data (i.e., data collected during study interviews and labelled with unique study code only). Neither RAs/field interviewers nor providers will have access to any archived research data (data entered into the database during a previous interview). Requests for data for analyses (e.g., by Research Team Members, graduate trainees and scientific team investigators) will be managed by the Data Manager and overseen by the PIs, the RC, and their senior delegates.

Data Storage:

Source documents are defined as original documents, data and records which may include evaluation checklists, videotaped observations and communication records (e.g., telephone logs, emails). Research staff will clearly define the various source documents used to support their study as part of their local data management process. Data collection will be completed by authorised study site personnel designated by the lead PI. Participants will not be identified in the study database by name or initials; they will only be identified by their unique participant ID.

Questionnaire data will be collected in electronic format using Qualtrics, an Application Service Provider (ASP) using a Software-as-a-Service (SaaS) platform for creating and distributing online surveys and other research services. Qualtrics uses Transport Layer Security (TLS) encryption for all transmitted internet data. Its services are hosted by trusted third party data centers that are audited using the industry standard SSAE-16 SOC Type 2 method. All data are stored within the region where data is collected. For our remote data collection procedures (during physical distancing public health measures due to COVID-19), individualized links will be sent to participants via email to the survey. If home visits resume at a time when it is deemed safe to do so, data will be collected on tablets that are password-protected and encrypted (using BitLocker).

Psychophysiological data will be collected through portable devices and transported in real-time to a password-protected and encrypted laptop; physiological data files are identified by the unique participant ID and not attached to any identifying information such as name or address. Any data collected on tablets or laptops will be uploaded to the secure server and encrypted data will be removed from the local drive. As highlighted in the section above, all data will be de-identified at the point of collection. De-identified data will be downloaded into a password-protected file on a secure McMaster server. Videotaped interactions (via Zoom or in person with camera) will also only be labelled with the participants' unique identifier and will be stored on the password protected server at McMaster.

Audio and video files of participant study visits as well as parenting program sessions are recorded through Zoom by RA's and group facilitators. RA's who conduct the visits save the files to the Zoom cloud after each visit with the participant ID and time point. Group Facilitators are instructed to save Zoom files with the agency name, group they are facilitating

(COSP or Triple P), and session number. Group session audio recordings, provider notes and consent to be contacted forms will all be shared and stored using password-protected files on secure servers accessible only to the Principal Investigators, RC and designates. Participant study intervention data and reports containing personal health information will be locked or securely stored online, separately from data containing study identification codes.

Audio files for transcription Using the following Zoom for Telehealth licensed requirements; PIPEDA_PHIPA Canadian Public Information Compliance Guide.pdf (zoom.us) qualitative interviews and focus groups are conducted and recorded by using Zoom Telehealth with consented participants. Zoom for Telehealth has technical security measures that comply with the Personal Information Protection and Electronic Documents Act (PIPEDA) and Personal Health Information Protection Act (PHIPA) to protect individuals' identities. The use of a third-party transcription service will be used. This third-party has a signed confidentiality agreement between themselves and the Principal Investigator. Audio files will be transferred to the transcriptionist using Sync.com which is a secure file transfer service using encrypted software. The focus group transcripts will be kept secure and private on a secure server (McMaster University). The transcripts will be destroyed 7 years following the final publication resulting from these studies.

Table 4: Types of Data and Storage Locations

Data Types	Storage Locations
Questionnaires (demographic data, caregiver and child wellbeing, dyadic/interpersonal relationship quality, parenting (style, attitudes, behaviour, mental and physical health, stress, childhood adversity, emotion regulation, therapeutic alliance, fidelity ratings)	
Audio Recordings (intervention group recordings, interviews, focus groups) and transcripts	
Random assignment and participant key (agency, group), contact information, and attendance tracking	
Computer-based tasks	
Video-taped parent-child interactions & coding	
Psychophysiological measures	
Integrative Clinical Evaluative Sciences (ICES) data linkage	

Data Back-Ups

Data back-ups are conducted monthly following McMaster Universities Research Data Management storage recommendations, which stipulate the following on a regular basis, performed by the DMT:

- 3 copies of all data stored
- 2 copies on hand on different systems (internal hard drive, external hard drive, cloud storage provider, etc)
- 1 copy in a separate location (“off-site”) from the others, with a trusted service provider.
    - E.g. PHF Data hosted in qualtrics.

https://rdm.mcmaster.ca/store#tab-backup-strategies

Data Quality Review

Audio and Video Home Visit Data:

- *Post-collection procedures*
- *Data pre-processing:.*

Group/Session Fidelity Audio, Video, and Survey Data:

- *Post-collection procedures:*
- *Data pre-processing:*
-

Psychophysiological Data:

- *Post-collection procedures:*
- *Data pre-processing:*

Questionnaire Data:

- *Post-collection procedures:*
- *Data pre-processing:*

Data Access Requests

How do I request access to PHF data?

External trainees and investigators may submit an application to request access to PHF data. The principal investigators will review the application and respond with feedback within 2 weeks of submission. Successful data requests will have provided clearly defined objectives within the scope of this study. The following will be required for the PHF data request application:

- Applicant and collaborator names, affiliation, and contact information
- Proposed project title
- Proposed project summary with defined objectives (500 word maximum)
- Variables and time points of interest (available in the PHF Data Dictionary upon request. Please contact the PI, Dr. Andrea Gonzalez, for access).
- Detailed analysis plan
- Proposed timeline
- Proposed outputs/products (e.g. manuscript submission to X journal).

Documentation and Metadata

What documentation and metadata accompany the data?

Any data that is shared will include the following. A README document describing the included Metadata will accompany all shared data with the following:

- Study background
- Research methods (how data was generated)
- Population characteristics
- Data Dictionary
- Analysis documentation

Data Dictionary (Screening, Baseline, Attendance, Post-program, 3-, 6-, 9- and 12 months etc.,) will be accompanied by a Data Dictionary spreadsheet, listing each item and possible response options. Scoring details for subscales, total scores, minimum and maximum scores, as well as reverse scoring is embedded. A README document explains the file names and protocols for arranging access to data included.

Audio files and transcripts (containing personally identified information such as signatures and full names); metadata associated with digital files (names of participants and interviewers); a spreadsheet of interview details and interviewee contact information; unpublished sample interviews for training purposes; and a README document that explains the restricted documentation, file names, and protocols for arranging access to restricted documents.

End-of-Project Planning

Data storage and destruction:

Succession Plan:

Data Linkage with ICES

Research data collected during the research interviews will be linked with participant personal data collected from health records. Participant consent will be obtained to access child personal data from electronic health records (ICES). ICES may also use personal health information under the authority of PHIPA s. 44 for approved research projects. With participants' informed consent, the PIs will share participants' unique identifiers (e.g., OHIP number), unique study code, and relevant study variables with ICES. Before the trial data is transferred to ICES, a data sharing agreement (DSA) between the principal investigator and ICES that governs the sharing of data will be established. The DSA governs the privacy and security of the information in the ICES data inventory. Most of the core health services data are governed under a DSA between ICES and the Ministry of Health and Long-term Care. ICES' policies, practices and procedures for using data are reviewed and approved on a regular basis by the Office of the Information Privacy Commissioner/Ontario. Data and the data dictionary will be transferred to ICES in a secure manner.

Data Safety Monitoring Board (DSMB)

This study will be overseen by a Data Safety Monitoring Board (DSMB). The RC will assist in the preparation of reports for the DSMB that include tables of all serious adverse events (AEs), enrollment and randomization figures, study withdrawals (by participant or investigator decision) and descriptions of participant flow. This evaluation will also assess data quality and timeliness, participant recruitment, accrual, and retention. AEs will be categorised as anticipated or unanticipated.

DATE	ATTENDEES	NOTES/DISCUSSION
		.
		.

Participant Confidentiality

Proper safeguarding techniques to protect the confidentiality of all data will be employed. All data will be de-identified prior to access, processing or analysis using a unique study code. Further safeguards to protect participant confidentiality will include limiting the collection of participant contact information to that which is necessary. Table 4 below outlines the justification for all identifying information that will be collected.

Table 5: Justification for Collection of Identifying Information

Identifier	Justification
Mailing address	Addresses are required for research interviews and to mail materials to participants
Email address	Email addresses are required to send zoom links, schedule interviews and groups, and to receive compensation.
Telephone Number(s)	Phone numbers are required for scheduling interviews and providing reminders based on participant requests. Further, obtaining multiple forms of contact information reduces missing data when participants move/are transient.
Personal health number - Ontario Health Insurance Plan (OHIP)	This information is needed to access data stored in ICES records
Full names	This information is needed for personal record data matches.
Date of Birth	This information is needed for demographic statistics
Sex	

Participants' identities will be protected throughout study. The 'key' that matches all participants' unique study codes to the full list of direct identities will be available only to the Principal Investigator and their delegates. This data will be kept in password protected files in encrypted files on secure servers or the computer of the PIs which are also password protected. All staff will be PHIPA trained. As part of the informed consent process, caregivers will be informed about privacy and confidentiality of data, and also about the potential need to breach confidentiality if concerns about any child's safety or imminent harm to any adult necessitate advising appropriate authorities (e.g., child protection, family physician, emergency medical services).

Staff Training & PHF Best Practices

- Team members are required to have secure McMaster University identification at the time of hire in addition to attesting to a Pledge of Confidentiality & Responsibility (see below).
- All communication between the team and participants occurs through these protected accounts. All data is locked using a secure password and is only shared between trained members of the research team. Members of the team receive training on how to access and handle password-protected files at the time of hire. These training procedures are outlined in the new employee on-boarding document.
- PHF team members participate in training that includes best practices for the use of strong passwords and two-factor authentication.
- Training documents outlining procedures for data storing and handling as well as confidentiality are reviewed with all new team members. Documents are updated as the study evolves, and team members maintain access to these documents via our team training resources hub (SharePoint) and refer to them frequently as well as receive training on any new processes that develop throughout the study.
    - Training includes meeting with a RC or senior member of the team to learn and review tasks, as well as the obligation to read protocols and standard operating procedures (SOP).
    - Mock visits may also be set up to practise skills as well as having new RAs observe seniors RAs conduct visits.
    - Reminders are sent from team leadership via email and Microsoft teams periodically to prevent drift.
- RAs handling any participant data/information must sign a confidentiality agreement (see below) as well as complete both the CITI Canada Privacy course and the Research Personnel (Privacy) courses.

RA Pledge of Confidentiality and Responsibility

I, _____, am participating in this research project as a research assistant. As such, I will have access to data collected during the course of the project. I hereby agree:

1. To keep all information obtained during the performance of my duties confidential (written, verbal, email, online, or other form). This includes, but is not limited to, all information about research participants, their families, their associates, and the organizations of which they are members, volunteers, or employees, as well as any information otherwise masked or known to be confidential.

2. To use the confidential information only for the purposes of the research project and not for any other purpose unless authorized to do so in writing by the co-Principal Investigators.

3. To return all confidential information provided to me in any form to the co-Principal Investigators when I have completed my tasks. If the confidential information is not returnable, I will erase or destroy it, including, without limitation, information stored on a computer hard drive or on a USB data storage device.

4. These obligations of confidentiality will continue after my participation in the research project has ended.

I understand that any unauthorized release or carelessness in the handling of this confidential information is considered a breach of the duty to maintain confidentiality.

I further understand that any breach of the duty to maintain confidentiality could be grounds for immediate dismissal and/or personal liability in any legal action arising from such a breach.

I also understand that failure to use and respect my time as a research assistant in the expected manner could be grounds for immediate dismissal and/or personal liability in any legal action arising from such a breach.

_____ _____

Signature of Research Assistant Date

_____ _____

Signature of Investigator Date

Qualtrics Delegation

Only members of the DMT (e.g., PI and RCs) may delegate roles and have access to participant questionnaire data on Qualtrics and our secure virtual private network. Senior RAs with the assigned role of participant visits and recruitment have access to participant contact information such as names, phone numbers, e-mails and mailing addresses for the purpose of contacting, scheduling, and conducting study visits. Training and ongoing supervision for RA's data handling is provided by RC at the time of hire and throughout the RA's involvement with the study.

Figure 2: Qualtrics Administrator Permissions (Principal Investigator & Research Coordinator)

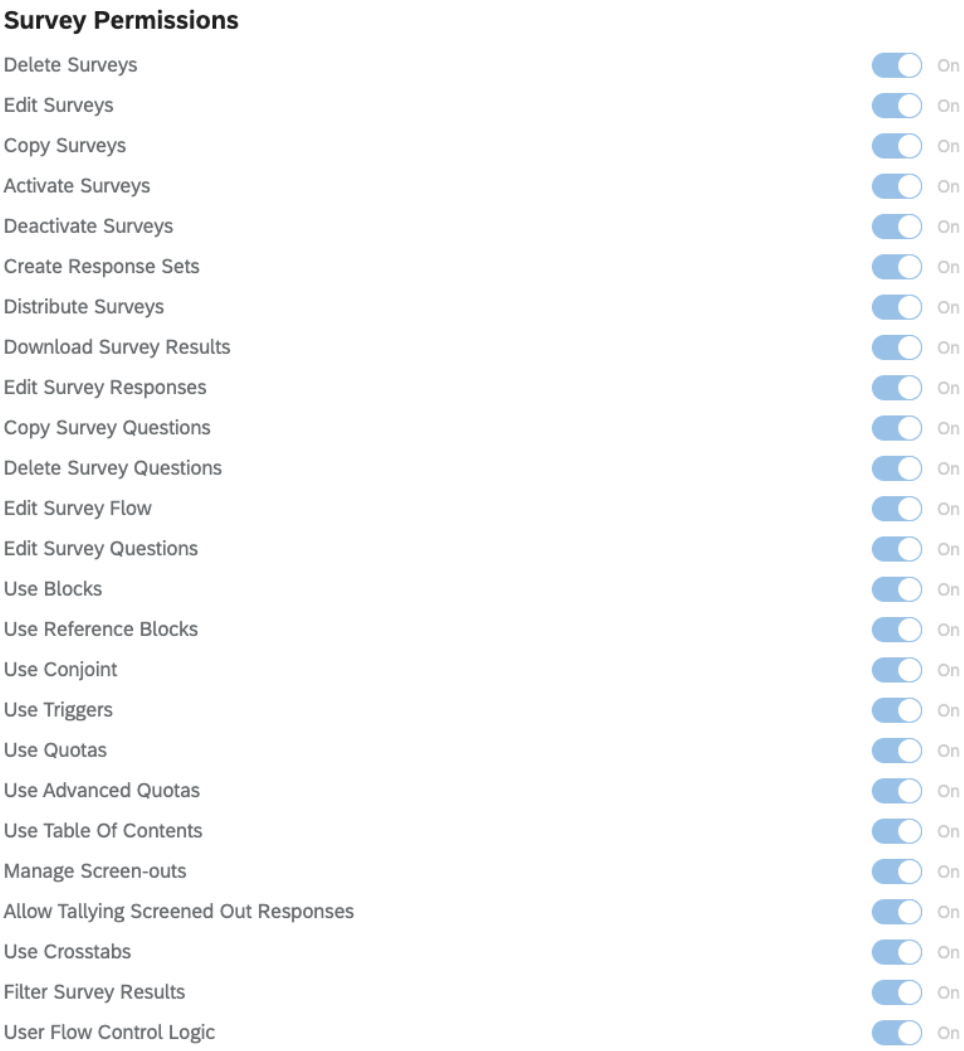
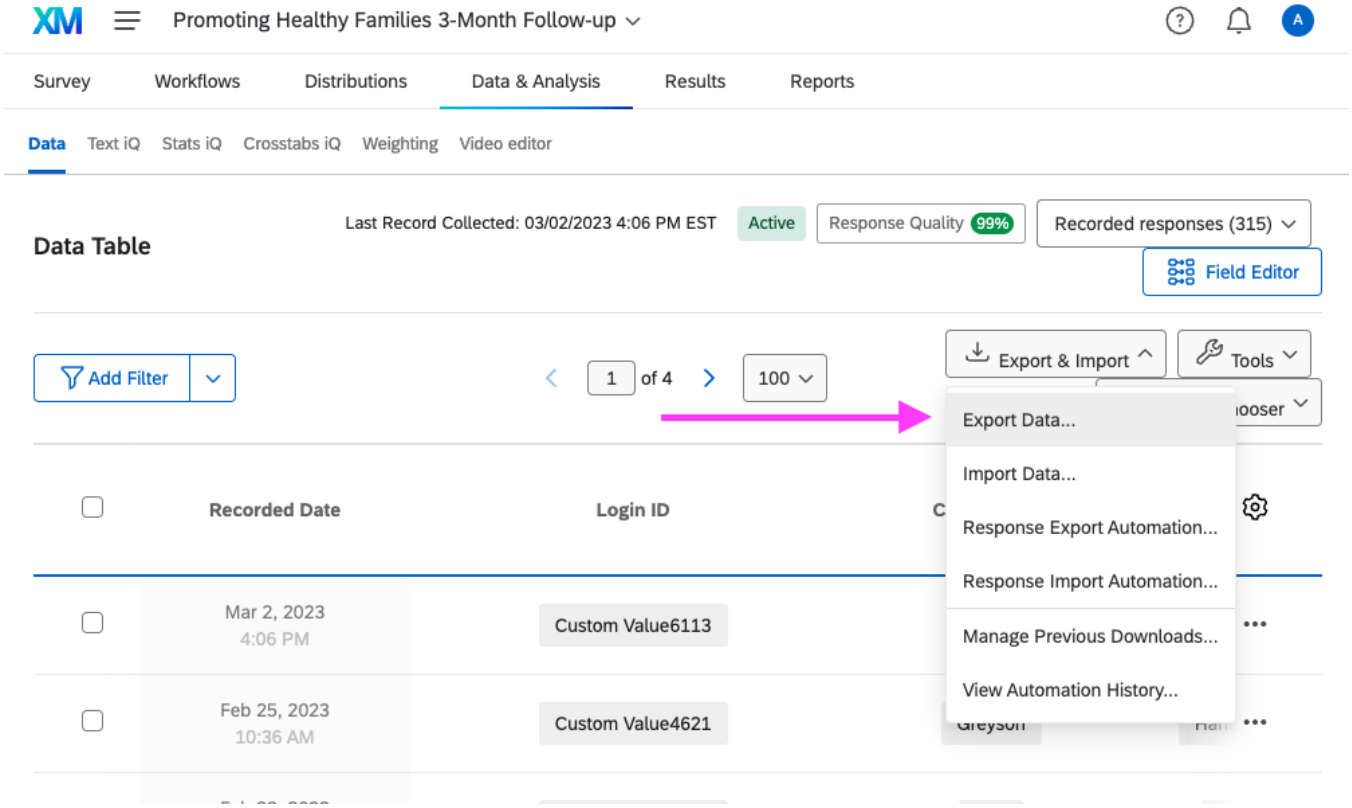


Figure 4: Qualtrics Raw Data Export

- Under the 'Data & Analysis' Tab for a survey, click the 'Export and Import' drop-down menu (see screenshot below), and then click 'Export Data…' for detailed export options.
- Saved on our secure drive (O:)
- Original files conserved and duplicated for analysis purposes
- De-identified data uploaded into statistical software (R, SPSS, nVivo, SAS, MPLUS, Excel) for the purposes of analysis.



DMP Authoring Rights

The DMP is a live working document that will be revised regularly by the team as the project progresses. The DMP will be reviewed annually by the Data Management Team.

Appendices

Administration of Measures

a. Time Requested of Participants

Participants randomized into one of the two parenting program intervention arms will attend eight sessions, each lasting two hours in length. Participants randomized into the treatment as usual (TAU) arm will receive another program or short-term therapy sessions, depending on the agency. All participants will also receive four study interviews conducted over the course of 14 months by a research study team member (please see table below). Interviews range in length from 90-120 minutes. Check-in calls will also be made at 3- and 9-month follow-ups. The total time requested to collect research data is no more than 10 hours. See Table below for schedule of measure administration.

b. Screening Measures

Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001). The SDQ is a behavioural screening questionnaire for 2 to 16-year olds. The questionnaire consists of 25 questions divided into five scales (see Appendix E). The prosocial behaviour scale showcases strengths, while the remaining four evaluate negative behaviours such as emotional symptoms, conduct problems, hyperactivity and inattention, and peer relationship problems. The SDQ has satisfactory reliability with mean internal consistency of Cronbach's alpha of 0.71 across scales for parents, mean cross-informant (parents and teachers) correlation of 0.37 across scales, and 3 to 6 month test-retest reliability of 0.62 (Goodman et al., 2001). The SDQ has concurrent validity with the Child Behaviour Checklist (Achenbach et al., 2000). Time to administer is approximately 5-10 minutes.

Parental Distress: The K6 scale of psychological distress (Kessler et al., 2003). The K6 is a brief, 6-item scale that provides a global measure of distress (drawing from depressive and anxiety-related symptomatology). Respondents are asked to report how frequently they experience six symptoms of psychological distress in the past 30 days (see Appendix E). The scale has a sensitivity of 0.36 and specificity of 0.96 against serious mental illness (Kessler et al., 2003), (and internal consistency of $\alpha = 0.89$ (Kessler et al., 2003). Time to administer is approximately 2-3 minutes.

Family Sociodemographic characteristics: Caregivers will be asked a few demographic questions including child age, parental age, marital status, and whether anyone in the household is currently receiving social assistance. Time to administer is approximately 2 minutes.

c. Primary Outcome Measures

Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001). The SDQ is a behavioural screening questionnaire for 2 to 16-year olds. The questionnaire consists of 25 questions divided into five scales. The prosocial behaviour scale showcases strengths, while the remaining four evaluate negative behaviours such as emotional symptoms, conduct problems, hyperactivity and inattention, and peer relationship problems. The SDQ has satisfactory reliability with mean internal consistency of Cronbach's alpha of 0.71 across scales for parents, mean cross-informant (parents and teachers) correlation of 0.37 across scales, and 3 to 6 month test-retest reliability of 0.62 (Goodman et al., 2001). The SDQ has concurrent validity with the Child Behaviour Checklist (Achenbach et al., 2000). We will administer the SDQ at all time points (baseline, immediately post-treatment, and 6- and 12-month follow-up), see Table 1.

Parenting Survey
The Parenting Scale (PS; Arnold, O'Leary, Wolff & Acker, 1993) is a measure of dysfunctional discipline practices in parents and is accepted internationally as a measure of parenting behaviour (Kliem et al., 2019). The 30-item questionnaire, has each item rated on a 7-point scale, measures three stable factors of dysfunctional discipline, including Laxness (LX: permissive discipline), Overreactivity (OR: authoritarian discipline, displays of anger and irritability), and Verbosity (VB: overly long reprimands or reliance on talking) (Tully et al., 2017; Arnold et al., 1993). The scale has adequate internal consistency with Cronbach $\alpha = 0.83$ for LX, $\alpha = 0.82$ for OR and $\alpha = 0.63$ for VB and good test-retest reliability with 0.83 for LX, 0.82 for OR and 0.79 for LX (Arnold et al.,1993; Tully et al., 2017)). For this study, we will be using the Overreactivity Scale (OS) as a primary indicator and the Laxness Scale (LS) as an additional measure; we will not administer the Verbosity Scale (VS;see Appendix E). The scale will be administered at all time points.

d. Secondary Outcome Measures

Observed Parenting (baseline; post-intervention; and 6-months)
Parenting behaviours will also be assessed using behavioural observations. Under COVID-19 protocols, at baseline and post-intervention using a recorded Zoom call, the parent and child will be observed during brief unstructured and semi-structured play interactions (5 minutes) and on a divided attention task where caregivers are asked to respond to questions from the interviewer while still attending to the child (5 mins).

At the 12-month follow-ups, if we will be allowed to conduct home visits, we will videotape caregiver child interactions in-person; if COVID physical distancing protocols continue we will adapt for Zoom protocols. At the 12-month follow-up, in-person or through Zoom, we will videotape caregiver-child interactions using three scenarios that resemble everyday parenting situations. The first involves a 10-minute teaching task – the dyad will be introduced to a puzzle

(Smart Game Castle Logix for 3- to 5-year-old children, and RoWood 3D Dolphin puzzle or STEM Children's Assembling Toy for 6- to 7-year-old children) designed to be too difficult for the child to complete alone. The parent will be instructed to help the child complete the puzzle through verbal assistance without doing the puzzle for the child. After the teaching task, the RA will acknowledge their effort and progress and collect the puzzle. The second task involves 5 minutes' free play with a standardized set of age-appropriate toys provided by the RA. The free play is in place to buffer potential stress generated during the challenging teaching task. Lastly, a clean-up task – at the end of the free play, the parent will receive instruction that once signaled by the RA, they should ask the child to clean up the toys into a basket. The clean-up task will last until the child finishes cleaning up the toys or, if the child refuses to clean up, a maximum of 10 minutes. Videotapes of the tasks will be coded using the Emotional Availability Scale (EAS; Birigen et al., 2000) which measures several aspects of parenting (e.g., sensitivity, involvement) across the tasks (see Appendix E). This coding scheme has been previously used by team members.

Integrative Clinical Evaluative Sciences (ICES) data linkage

The ICES data repository consists of record-level, de-sensitized and linkable datasets. Data for the secondary outcomes will be derived from various ICES databases including Health Services Administrative Data (physician billings, inpatient hospital discharges and emergency and ambulatory care visits). Specifically, our secondary outcomes of interest are related to maltreatment-related injuries or hospitalizations and include the following data: (1) Maltreatment-syndrome (ICD9 codes 995.5, E967, 994.2 or 994.3 and ICD10 codes T74, Y06, Y07, or T73, T740, T743, T748, T749); (2) Assault (ICD9 codes E960-E969, and ICD10 codes X8-Y09, X85-Y05, Y10-Y34, Y871, R456, T741, T742, Z045, Z616, and K018 and K021 (OHIP DXcode)); (3) Undetermined cause (ICD9 codes E980-989, V68.2, V70.4, V71.4, V71.5, V71.6, V71.8, and V71.9 and ICD10 codes Y10 -Y34, Z040, Z0450, Z0451, Z0458, Z048); (4) Adverse social circumstances (ICD9 codes V15.4, V15.5, V15.9, V60 (V600-V619) and ICD10 codes Z60 - Z63, Z72, Z74, Z76.1, Z76.2, Z81, Z86.5, Z91.6, Z91.8); and (5) Social, Marital or Family Problems (ICD8 897-899, 900-907, 919, 949, 959). The participant consent form will include permission to collect personal health card information for the purposes of transferring trial data to ICES for linkage (please see participant informed consent form). Before the trial data is transferred to ICES at 12-months follow-up, an agreement between the Principal Investigator (Andrea Gonzalez) and the lead ICES investigator (Astrid Guttman) that governs the sharing of data will be established. It is important to note the following: 1) all linkages are anonymous within ICES; 2) the agreement is a data transfer to ICES; 3) ICES is a prescribed entity under Privacy Health Information Protection Act (PHIPA) compliant; and 4) only summary data will be provided to the research team. Accessing personal record data has been chosen for two reasons: first, to reduce participant burden; second, to compliment caregiver self-report data and observational data; and third, to permit longitudinal follow-up on key measures. The ICES data will be also used for an economic evaluation (see section 4.4).

e. Exploratory Measures

Composite Caregiving Questionnaire: The CCQ was developed to incorporate a number of parenting features including empathy, caregiving reflection and mentalization (Maxwell et al., 2020). It consists of 42 items incorporating: parental self-efficacy (12 items), hostile parenting (5 items), caregiving helplessness (7 items), and maternal reflective functioning (18 items). The CCQ takes approximately 10 minutes to complete and will be administered at all time points (see Appendix F).

Child Version, Parent-Child Conflict Tactics Scale - Revised (CTSPC-R): This CTSPC-R is a picture-based version of the original Parent-Child Conflict Tactics Scale (Straus et al., 1998) for children aged 4-16 years of age capturing acts of non-violent discipline, and psychological and physical abuse towards the child in the last year. We only use 4 items It is comprised of 22 items, with items rated on a five-point scale (0 = did not occur, 1 = did occur once, 2 = did occur a few times, 3 = many times, 4 = every time). Companion picture cards that have been developed and validated for younger aged children, 3-8 years (Sierau, et al., 2018). In addition, for the younger age group a visual representation of the five-point scale has been developed to allow for pointing instead of providing a verbal answer (Kantor et al., 2004). A recent study examining the psychometric properties of the measure found a three-factor model corresponding to the severity of the questions. For the purposes of this study only Module 1 (minor severity, 4 items) comprised of incidences of nonviolent discipline (time-outs, distraction, reasoning and removal of toy)will be administered (please see Appendix F). Administration of the 4 items takes approximately 5 minutes to complete. This test will only be administered at the 12-month follow-up and only if in-person visits resume.

The Difficulties in Emotion Regulation Scale – Short Form (DERS-SF): The DERS-SF is an 18-item scale measure used to assess deficits in regulating emotions. The DERS-SF is derived from the 36-item DERS, and has excellent psychometric properties (correlation between the DERS and the DERS-SF ranges from .91 to .98) (Kaufman et al., 2016) A total score is provided, as well as scores on 6 subscales. The DERS-SF takes approximately 3 minutes to complete and will be administered at all time points (see Appendix F).

Psychophysiological Assessment of Parental Emotion Regulation and Parent-Child Coregulation: At the 12-month follow-up, if the visit occurs in-person, we will collect ECG and EDA; measured as skin conductance data from parent-child dyads during the three sections of interaction. Physiological data will be collected from the subsample participating using the wearable BIOPAC BioNomadix devices (BIOPAC Systems, Inc., Goleta, CA, USA). These portable, non-invasive devices will be connected to a set of adult/child electrodes attached to the surface of each participant's skin (one on the right clavicle, two on the lower left and right rib, and two on the palm of the non-dominant hand). ECG and EDA signals will be recorded in real-time during the caregiver-child interaction and transported wirelessly to a receiver connected to a laptop with no internet connection. Prior to putting on the electrodes, the research assistant

will explain the device and procedures to the dyad, making sure they are comfortable with the assessment. After putting on the electrodes, the dyad will be given time to adjust before starting the interactions. The BIOPAC system has been widely applied to measuring peripheral physiology including ECG and EDA among parents and young children in the current age range (e.g., Song, Colasante, & Malti, 2018; Woody et al., 2016). Additionally, this psychophysiological assessment protocol has been used by our team members in previous work (e.g., Ravindran, Zhang, et al., 2021; Zhang, Han, & Gatzke-Kopp, 2021).

In addition to collecting the physiological measures, the videotapes of the two challenging sections of the interaction (the teaching task and the clean-up task) will be coded by team members on second-by-second challenging child behaviours (e.g., non-compliance, disruptive behaviours) and parental sensitivity. The Challenging Child Behaviours Micro-Coding Scale is adapted from work by Lorber and O'Leary (2005; see Appendix F. Parental sensitivity will be coded by adapting the sensitivity subscale in the Emotional Availability Scale into a micro-coding scheme. That is, we will adopt the same operationalization of sensitive/insensitive behaviours; however, instead of giving a global rating of how sensitive parenting behaviours are during the entire interaction, discrete behaviours will be identified, rated, and mapped temporally to obtain time-series ratings of parental sensitivity.

From the ECG data, we will calculate parents' and children's moment-to-moment cardiac arousal (measured as inter-beat intervals; IBI) and estimates of heart rate variability (HRV) at specific frequency ranges, reflecting parasympathetic inputs to cardiac activity. From the EDA data, we will calculate parents' and children's moment-to-moment skin conductance level (SCL), reflecting sympathetic arousal. Together, these measures will enable examination of their dynamic physiological reactivity during the interaction and how such reactivity is associated with observational measures (e.g., child and parent behaviours).

Based on the physiological and behavioural measures, parent emotion regulation during the interactions will be examined as a set of dynamic processes capturing how parental physiology responds to children's challenging behaviours, and the relations between parents' physiological and behavioural responses. Meanwhile, parent-child coregulation will be examined as the dynamic concordance between their physiological arousal, and how such concordance is moderated by the dynamics in sensitive parenting behaviours.

Parenting Daily Hassles Scale: This scale is designed to measure parental perceptions about the minor daily hassles and inconveniences associated with parenting. It comprises 20 items that describe discrete events that involve challenging child behaviour or various tasks associated with parenting (Crnic & Greenberg, 1990). Both frequency and intensity scores are obtained. Good internal consistency and excellent convergent validity have been reported. The Daily Hassles Scale takes approximately 7 minutes to complete and will be administered at all time points (see Appendix F).

f. Parental Attributions

The Parent Cognition Scale: The Parent Cognition Scale is a 16-item self-report measure designed to assess the degree to which parents endorse dysfunctional child-responsible and parent-causal attributions for child misbehaviour. All items were taken from actual parent attributions from mothers of toddler and preschool children, recorded during a previous study (Slep & O'Leary, 1998). Parents think about their child's misbehaviour over the past 2 months, and rate various possible causes for this behaviour (see Appendix F). Response options are rated on a 6-point Likert scale ranging from 1 (always true) to 6 (never true). Demonstrates adequate internal consistent (αs .81–.90), test-retest reliability (rs .55–.76), and convergent and discriminant validity. This questionnaire will be administered at baseline, post-treatment, and 6-and 12-month follow-up. It takes approximately 5 minutes to complete.

Plotkin Child Vignettes (PCV; Plotkin, 1983): The PCV assesses participants' judgements of the intentionality of child misbehaviour in 18 vignettes. Participants indicate how much they consider the child tried to intentionally annoy on a 9-point scale, from (1) did not mean to annoy me at all to (9) the only reason the child did this was to annoy me. Item scores are summed for a total score wherein higher scores indicate more negative child behaviour attributions. Validity has been demonstrated through associations with an analog assessment of attributions (Rodriguez, Cook, & Jedrziewski, 2012) and abusive mothers attain higher scores relative to comparison parents (Haskett et al., 2006; Plotkin, 1983). Administration takes approximately 10 minutes. This test will be administered at baseline and 6-month follow-up.

g. Other Measures

Demographics: During the baseline assessment demographic questions will be asked to collect the following information: household composition, ethnicity of caregiver(s) and child, employment status of caregiver(s), caregiver(s) and child country of birth, language(s) spoken in the home, caregiver(s) and child health chronic health condition, perception of caregiver(s) and child health, gestational age of child at birth, mode of delivery, and special care required at birth (e.g., NICU) (see Appendix G).

Caregiver history of Childhood Adversity: We will administer the Childhood Trauma Questionnaire (CTQ; Bernstein & Fink, 1998) and select items from the Adverse Childhood Experiences survey (ACES; Felitti et al., 1998) for identification of childhood history of maltreatment and other adverse experiences. The CTQ is a 28-item brief self-report questionnaire that retrospectively assesses five types of abuse, history of sexual, physical, and emotional abuse, and physical and emotional neglect. Subjects rate statements about lifetime childhood experiences on a five-point scale ranging from "never true" to "very often true". Reliability and validity of the CTQ, including stability over time, convergent and discriminant validity with structured trauma interviews and corroboration using independent data have been determined (Bernstein et al., 2004; 1994; Scher et al., 2004). It has demonstrated high

internal reliability, with Cronbach's alphas ranging from 0.74 to 0.90 and good test-retest reliability at 3-months. Factor analyses tests on the five-factor CTQ model showed structural invariance which demonstrates good validity. We will administer four items from the ACEs survey regarding childhood history of parental mental health, criminality, and parental loss/separation or divorce. These questionnaires will only be administered at baseline (see Appendix G).

h. Caregiver Mental Health

The Center for Epidemiologic Studies Depression scale (CES-D): The CES-D is a short structured 20 item self-report scale designed to measure current level of depressive symptomatology in the general population (Radloff et al., 1977). The scale asks about symptoms that occurred in the week prior to the interview, with items (frequency of symptoms) rated on a 4-point Likert scale. The scale takes between 2 and 5 minutes to complete (Vilagut et al., 2016). The CES-D is one of the most widespread brief scales for measuring depressive symptoms (Vilagut et al., 2016) and shows high internal consistency, acceptable test-retest stability, excellent concurrent validity by clinical and self-report criteria, and substantial evidence of construct validity (Radloff et al., 1977). The CES-D provides cut-off scores (e.g., 16 or greater) that help to identify those at risk for clinical depression, with good sensitivity and specificity and high internal consistency (Lewinsohn et al., 1997). The CES-D will be administered at all time points – baseline, post-intervention, 6- and 12-months points (see Appendix G).

Generalized Anxiety Disorder (GAD-7): The GAD-7 (Generalized Anxiety Disorder – 7) (Spitzer et al. 2006) is a brief, 7-question scale used to screen for presence and severity of Generalized Anxiety Disorder (GAD) (Spitzer et al., 2006). The length of time to administer is 1-2 minutes (Mossman et al., 2017). The GAD-7 has demonstrated good psychometric properties, including internal consistency (Cronbach $\alpha = 0.92$) and convergent validity with the Beck Anxiety Inventory ($r = 0.72$) (Spitzer et al., 2006). The GAD-7 will be administered at all time points (see Appendix G).

Alcohol and Substance Use: Alcohol Use Disorders Identification Test Consumption (AUDIT-C): The AUDIT-C is a three-item alcohol screen derived from the World Health Organization's AUDIT to determine hazardous drinking or active alcohol use disorders (including alcohol abuse or dependence). It is scored on a scale of 0-12 (scores of 0 reflect no alcohol use). In men, a score of 4 or more is considered positive; in women, a score of 3 or more is considered positive. Generally, the higher the AUDIT-C score, the more likely it is that the patient's drinking is affecting his/her health and safety. The AUDIT-C performs as well as the AUDIT in a primary care population (Bradley et al., 2007) and performed well in the U.S. general population (Dawson et al., 2005). In addition to the AUDIT-C, we will ask one question about medication use and three questions about cannabis use (please see Appendix G). These questions take two minutes to complete and will be asked at all time points.

i. Interpersonal Conflict/Relationships

Abbreviated Dyadic Adjustment Scale (DAS-7; Sharpley & Rogers, 1984): The DAS-7 is a 7-item measure designed to assess the relationship quality of intact couples. This shortened version of the original DAS-32 includes items aimed at assessing relationship satisfaction and the degree to which the couple agrees on matters of importance to the relationship. The DAS-7 takes 2 minutes to complete and will be administered at all time points (see Appendix X).

Marital Conflict (Timmons et al., 2017): The marital conflict questionnaire consists of 12 items assessing daily marital conflict. Scores per item ranged from 0 (not at all) to 3 (a lot). Six items assessed the actions of the reporter (i.e., I was angry at my spouse; I was annoyed with my partner/spouse; I yelled at or criticized my partner/spouse; I felt distant or withdrawn from my partner/spouse; I nagged my partner/spouse; I flew off the handle or exploded at my partner/spouse). An identical six items assessed the actions of the spouse (My partner was angry at me; My partner/spouse was annoyed at me, etc.). In a previous study (Timmons et al., 2017), wife and husband reports were significantly correlated ($r = .48$, $p < .001$), which is to be anticipated given the dyadic nature of conflict. Caregivers will complete the questionnaire at all time points (see Appendix G).

Therapeutic Alliance Assessments: During the group sessions (Triple P or COS-P), or individual sessions (Triple P sessions 5-7 and TAU), we will collect bi-weekly assessment of alliance using the Working Alliance Inventory (WAI; group – short form) or Working Alliance Inventory (individual sessions – short form), please see Appendix G. The WAI assesses three key aspects of therapeutic alliance: i) agreement on tasks of therapy, ii) agreement on goals of therapy, iii) development of an affective bond. The Working Alliance Inventory-Short Revised (WAI-SR) demonstrated good psychometric properties in validation studies with outpatients and inpatients (Munder et al., 2010). We will only administer items from the Goals and Bond subscales (total 8 items). Item responses range from 1= 'Seldom' to 5= 'Always'.

j. Economic Evaluation Measures

We will administer three questionnaires for the economic evaluation across all time points. The Child and Adolescent Use Schedule (CA-SUS) – a modified version will be used to assess child health and social care service use at baseline, 3-, 6-, 9- and 12-months (Byford et al., 2007). We will also use the Health Utility Index – Mark 3 (HUI3) for both children (parent-proxy) and for caregivers (self-report) which captures eight domains (vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain/discomfort), with 1 to 2 items per domain and 4 to 6 response alternatives per item (Furlong et al., 1992). Together, the 12 items generate an overall health-related quality of life (HRQOL) utility score. The HUI3 has demonstrated good discriminant validity and high test-retest reliability (Boyle et al., 1995). We will also use the Health Status Classification System Preschool Version (HSCS-PS; Saigal et al., 2005). This 14-item HS instrument assesses the following 12 health attributes: seeing, hearing, speaking, getting

around, using hands and fingers, taking care of self, feelings, learning and remembering, thinking and solving problems, pain and discomfort, general health, and behaviour, each with 3 to 5 levels of severity. Scores for the 12 health attributes of the HSCS-PS were dichotomized into either “no health problem” or “any health problem,” which could be mild, moderate, or severe.

Timing of Measures

Table 6: Timing of Measures

Variable	Measure	Baseline	Post-program	6M Follow-up	12M Follow-up
Primary Outcome					
Child behavioural problems	Strengths and Difficulties Questionnaire (SDQ; 2-4 years; 4-18 years)	X	X	X	X
Parental harsh discipline	Overreactivity Scale	X	X	X	X
Observed parenting (parental sensitivity)	Emotional Availability Scale (EAS)	X	X	X	X
Secondary Outcomes					
Average number of maltreatment-related child injuries and hospitalization and family problems	(1) Maltreatment-syndrome (ICD9 codes 995.5, E967, 994.2 or 994.3 and ICD10 codes T74, Y06, Y07, or T73, T740, T743, T748, T749); (2) Assault (ICD9 codes E960-E969, and ICD10 codes X8-Y09, X85-Y05, Y10-Y34, Y871, R456, T741, T742, Z045, Z616, and K018 and K021 (OHIP DXcode)); (3) Undetermined cause (ICD9 codes E980-989, V68.2, V70.4, V71.4, V71.5, V71.6, V71.8, and V71.9 and ICD10 codes Y10 -Y34, Z040, Z0450, Z0451, Z0458, Z048); (4) Adverse social circumstances (ICD9 codes V15.4, V15.5, V15.9, V60 (V600-V619) and ICD10 codes Z60 - Z63, Z72, Z74, Z76.1, Z76.2, Z81, Z86.5, Z91.6, Z91.8); and (5) Social, Marital or Family Problems (ICD8 897-899, 900-907, 919, 949, 959).				X
Additional Outcomes					
Parenting capacity					
Parental stress	Parenting Daily Hassles	X		X	X
Self-efficacy, empathy, helplessness, and reflective function	Composite Caregiving Questionnaire	X	X	X	X

Emotion regulation	Difficulties in Emotion Regulation Scale (DERS-SF)	X		X	X
Parental attributions	Parent Cognition Scale	X	X	X	X
Parental attributions	Plotkin Child Vignettes (PCV)	X		X	
Child outcomes					
Parenting practices	Parent-Child Conflict-Tactics Scale-Revised (CTSPC-R)				X
Responsiveness	Emotional Availability Scale (child scales)	X	X	X	X
Other					
Parental alcohol and substance use	AUDIT and medication and cannabis use questions	X		X	X
Parental mental health	Centre for Epidemiological Depression Scale (CES-D-10); Generalized Anxiety Disorder (GAD-7)	X	X	X	X
Parental history of childhood adversity	Childhood Trauma Questionnaire (CTQ); Adverse Childhood Experiences questions	X			
Partner conflict	Marital Conflict Questionnaire Dyadic Adjustment Scale (DAS-7)	X	X	X	X
Economic evaluation					
Health Service Utilization for Children (caregiver report)	Child and Adolescent Service Use Schedule (CA-SUS)*	X		X	X
Overall Quality of Life and Health Measure for Child	HUI3 proxy parental report and Health Status Classification System Preschool Version*	X		X	X
Quality of Life Measure for Caregiver	HUI3*	X		X	X
Health Care Utilization	Data from ICES: Canadian Institute for Health Information (CIHI) Hospital Discharge Abstract Database (DAD, for acute inpatient care), National Ambulatory Care Reporting System (NACRS, for emergency department visits and selected outpatient clinic visits) and Ontario Health Insurance Plan (OHIP) Claims History Database.[t]				X

Note: * For the economic evaluation, these measures will also be administered at the 3- and 9-month time points. [t] ICES data will also be requested at study end for one-year data prior to baseline and 1-year data after baseline.

Economic Evaluation

A 12-month trial-based economic evaluation will compare the costs and Quality Adjusted Life years (QALYs) associated with Triple P versus TAU and COSP versus TAU from a societal perspective. In our primary analysis, the costs considered in the analyses include the costs associated with the development and implementation of the interventions (Triple P, COSP) which will be collected as part of the trial. In addition, child-related healthcare resource use consumed by each group (e.g. hospitalizations, emergency room visits, physician and other healthcare professionals visits) will be captured in the trial using the CA-SUS questionnaire which will be administered to the caregiver. Healthcare resource utilization derived from the CA-SUS will be multiplied by their respective unit costs using public data from Ontario. Once the study data will be transferred to ICES (see section 6.2.3), we will also conduct an additional costing analysis using the healthcare resources and costing information contained in the administrative data. Specifically, the records of all children who participated in the study will be linked to several key databases housed at ICES, including the Canadian Institute for Health Information (CIHI) Hospital Discharge Abstract Database (DAD, for acute inpatient care), National Ambulatory Care Reporting System (NACRS; for emergency department visits and selected outpatient clinic visits) and Ontario Health Insurance Plan (OHIP) Claims History Database. One year of administrative data before and after study enrolment will be monitored to document healthcare resource utilization. Healthcare expenditures will be determined using standardized ICES costing algorithms (Wodchis et al., 2013).

To calculate the QALYs associated with each intervention, our primary analysis will use the HUI3 (caregiver proxy report). An area under the curve approach will be used to calculate the QALYs by weighting the utility scores derived from the HUI3 by time spent in health state. In a secondary analysis we will use the Health Status Classification System Preschool Version (HSCSPV) to calculate the QALYs. Bootstrap techniques (Efron et al., 1993) will be used to deal with sampling uncertainty and to generate confidence intervals around costs and QALYs. Cost-effectiveness acceptability curves will be used to summarize the uncertainty and to present the probability that Triple P or COSP is cost-effective compared to usual care at certain cost-effectiveness willingness to pay threshold (e.g. $50,000 or $100,000 per QALY gained). The net-benefit regression framework (Hoch et al., 2002; 2006) may also be used to adjust for any imbalances (e.g. quality of life) between the groups at randomization. The impact of key variables on the results (e.g. cost of Triple P or COSP) will be explored through sensitivity analyses. Missing data will be imputed using multiple imputations (Carpenter et al., 2013). The economic evaluation will be conducted in accordance of Canadian and international guidelines for the conduct of economic evaluations of healthcare programs (Canadian Agency for Drugs and Technologies in Health (2017); Drummond et al., 2015; Ramsey et al., 2015), The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) guidelines (Husereau et al., 2013) will be used to report the results.

Statistical Analyses

Baseline Comparisons: Descriptive statistics (means, standard deviations, relevant quantiles, and proportions) will be used to compare study arms with respect to measures taken at baseline, including caregivers and children characteristics such as demographics.

Analyses of Effectiveness Outcomes: All analyses will adhere to the intent-to-treat principle. The primary analysis will consist of two pair-wise comparisons: Triple P vs TAU, and COSP vs TAU. The level to declare significance will be 5%, two-sided. Within each pair-wise comparison Bonferroni correction will be used to limit the type I error probability to 5% for the three primary outcomes, and to limit the type I error probability to 5% for the five secondary outcomes. No correction will be used in the comparison of additional outcomes. Ninety-five per cent confidence intervals will be provided for parameter estimates. An exploratory analysis will compare Triple P vs COSP.

Primary Outcomes: A mixed-effects linear model will be used to compare study arms. The analysis will include measurements at all three post-randomization time-points. The fixed effects will be study arm, time-point, and the corresponding measure at baseline. There will be a random effect for subject. Overall study arm mean differences will be estimated. Study arm by time-point interaction will be added to the model to examine for temporal effects.

Secondary Outcomes: We expect very sparse data. Hence, for each of the five categories, each subject will be scored a “1” if they have one or more corresponding International Classification of Diseases (ICD) codes for that category, and a “0”, otherwise. A logistic regression will be used to compare study arms. Odds-ratios will be estimated.

Additional Outcomes: Except for the child measure of parenting practices (CTSPC-R), which is measured at month 12 only, the additional outcomes will be compared between study arms using a model described in 4.1.1. A two-sample t-test will be used to compare study arms with respect to CTSPC-R.

Subgroup Analyses: Since TAU will depend on site, a site subgroup analysis for the primary outcomes will be performed. Fixed effects for site and site-by-study arm will be added to the model described in 4.1.1. A similar subgroup analysis will be performed for the indicator variable: therapy primarily delivered online (yes=1; no=0).

Missing Data: Multiple imputation (Rubin, D.B. (1987), Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons, Inc.) will be used to account for missing values.

Sample Size: Separate type I error probabilities will be used for each pair-wise comparison (Triple P vs TAU, COSP vs TAU). Using a two-sided type I error probability of 0.05 (Bonferroni corrected for three

primary outcomes) and a 15% lost-to-follow-up, a total of 600 subjects (200 per study arm) will provide an 80% power to reject the null hypothesis of no treatment effect if, in each pair-wise comparison, the active arm is superior by 0.35 of a standard deviation. From previous trials (references) the estimates of the standard deviation for the Overactivity Scale (OS) lie between 0.75 and 1.0, with a mean at baseline of around 3.5. Thus, we would have adequate power for differences between 0.26 and 0.35 on the OS scale. Differences smaller than 0.35 would not be considered clinically relevant. From a previous trial (Spijkers et al., 2013) the estimate of the standard deviation for the SDQ is 4.5, with a mean at baseline of around 13.5. Thus, we would have adequate power for difference around 4.7 on the SDQ scale. Differences smaller than 4.7 would not be considered clinically relevant.