GenPMTO: Pilot study

Evaluation protocol

19/03/2025

Jack Martin,
Niall Daly,
Lilli Wagstaff,
Emma Forsyth,
Tom McBride.

Acknowledgements

We are grateful to colleagues from Barnardo's (including Daniel Chery, and Anna Fitzgerald) and colleagues from Implementation Sciences International (including Dr Abigail Gewirtz, Dr Marion Forgatch, and Laura Rains), for their assistance with and input into the evaluation. We are also grateful to colleagues from the Behavioural Insights Team (particularly Dr Patrick Taylor, and Leonie Nicks), the Youth Endowment Fund (particularly Dr Mollie Bourne, Dr Daniel Acquah, and Fardawza Ahmed), Dr Sajid Humayun (University of Greenwich) and Dr Kirsten Asmussen (Foundations) for their helpful advice and feedback.

We are grateful to the Youth Endowment Fund (originally funded by an endowment from the Home Office) and Stuart Roden for the funding which supports this work, and to the Behavioural Insights Team and Nesta for their support and role in administering this funding.

THE BEHAVIOURAL INSIGHTS TEAM

YOUTH ENDOWMENT FUND

nesta

About

This report was first published in March 2025, and is available to download as a free PDF at: https://www.bi.team/wp-content/uploads/2025/03/GenPMTO-Pilot-study-protocol.pdf

This work is being undertaken by the Ending Youth Violence Lab at the Behavioural Insights Team. The Behavioural Insights Team (BIT) is a global social purpose company that generates and applies behavioural insights to inform policy, improve public services and deliver results for citizens and society. With company number 08567792.

Contents

Project title	GenPMTO Evaluation - Pilot study
Developer (Institution)	Implementation Sciences International Inc. (ISII)
Delivery partner (Institution)	Barnardo's
Evaluator (Institution)	Ending Youth Violence Lab (part of the Behavioural Insights Team)
Principal investigator(s)	Tom McBride
Evaluation plan author(s)	Jack Martin, Niall Daly, Lilli Wagstaff, Emma Forsyth, Tom McBride
Evaluation setting	3 boroughs in London (Barking & Dagenham, Brent, and Tower Hamlets)
Target group	Caregivers of young people aged 8-14, who have been identified as being at risk of committing violent acts and/or entering into the criminal justice system.
Planned number of participants	Approximately: ● 100-120 primary caregivers (quantitative sample) ● 9 caregivers (qualitative sample) ● 6 practitioners (qualitative sample) ● 6 referrers (qualitative sample) ● 2 Barnardo's managers (qualitative sample)
Funding source and declaration of interests	Funding was received for this work, from the Youth Endowment Fund (YEF) and from philanthropist Stuart Roden (via Prism the Gift Fund). There are no known conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome.
Version history	Note that protocol publication was delayed from December '24 to March '25, as recruitment challenges delayed the start of the study. The protocol has been published prior to randomisation for the first parent cohort (originally expected to occur in December '24/January '25).

1. Study rationale and background

About the Ending Youth Violence Lab

The Ending Youth Violence Lab ('the Lab') was founded in Summer 2022, bringing together expertise in intervention, evaluation and youth violence. It is funded by Stuart Roden and the Youth Endowment Fund (YEF) and is being incubated at the Behavioural Insights Team (BIT).

The Lab's mission is to catalyse a step change in understanding and tackling youth violence. To do this, we do 3 things: Firstly, we identify promising interventions which seek to address youth violence. Secondly, we fund the development and delivery of these interventions. Thirdly, we conduct research to assess the delivery of interventions, identify ways to improve them, and explore the potential for further evaluation (with a focus on early-stage testing, to support the work of YEF).

We prioritise three strands of activity:

1. **Supporting the importation, adaptation, and testing of well-evidenced interventions from overseas -** we identify approaches with strong evidence of improving youth violence outcomes or related upstream factors in other countries, adapt these to the UK context, and deliver early-stage testing.

2. **Working with UK organisations to develop strong ideas into evaluable interventions -** we work with the youth violence prevention sector to find interventions that have strong theoretical underpinnings, are committed to rigorous evaluation, and oversee the development and early-stage testing needed to get them trial-ready.

3. **Working with developers, researchers, practitioners, and service users to co-design new and innovative approaches -** we build partnerships and fund the development of novel approaches to tackling youth violence, with a focus on addressing underserved populations and unmet needs.

The project described in this protocol forms part of strand 1 of the Lab's approach.

Project overview

About the intervention

This project is a multi-stage evaluation of GenPMTO (Generation Parent Management Training – Oregon Model). GenPMTO is a parenting programme which involves trained practitioners using active teaching approaches (such as group problem-solving, role-play, and video modelling) to support caregivers in using positive parenting strategies at home. The programme is designed to improve parenting practices, as well as a range of outcomes for young people, including improving academic performance, reducing school exclusions, and reducing offending and criminal behaviour.

The programme was originally developed in the USA, though now operates internationally, including in Denmark, Iceland, Mexico, Chile, Canada, the Netherlands and Norway, where it has been delivered and successfully evaluated. However, GenPMTO has never been delivered in the UK.

Evidence registries in the USA and UK classify the programme's effectiveness as being strongly supported by research evidence.[1,2] Importantly, and rarely for programmes of this type, GenPMTO has long-term evidence of reducing arrests.[3,4]

About the evaluation

This project represents the first attempt to deliver and evaluate GenPMTO in the UK. To do so, the Lab is conducting a multi-stage evaluation, involving delivering the intervention across three London boroughs, with a focus on caregivers of 8-14-year-old children and young people (CYPs), who are identified to have risk factors associated with involvement in violence.

- The first, completed, stages of the project are **Stage 1 & 2 - Adaptation, training and feasibility study**. Here, we tested the extent to which it is feasible to deliver and evaluate GenPMTO in the UK. By adaptation, we mean making a series of changes to the programme to attempt to maximise its acceptability and feasibility in a UK context. On the basis of positive findings (to be published at the end of the overall project), progression to the next stage was approved.
- The upcoming stage, described in this protocol is **Stage 3 - Pilot trial.** Here, the Lab will further test the extent to which it is feasible to robustly evaluate this programme, and gather preliminary evidence on the programme's impact during a pilot trial. If these results are positive, we will progress to the fourth stage.
- The final stage will be **Stage 4 - Efficacy trial.** Here, we will conduct an efficacy trial, to robustly determine if the programme can have a positive impact on outcomes for families and young people in the UK. This would involve extending delivery by a further 9-12 months, delivering the intervention to an additional cohort of families, and analysing their outcome data to make a robust assessment of the impact of GenPMTO. As part of this stage 4 study, the new outcome data collected in stage 4 would be analysed alongside the outcome data collected from the stage 3 pilot trial to provide an estimate of the effect of the programme. This is known as an efficacy trial with an 'internal pilot', and is a well-established approach in cases where a pilot trial is successful and does not indicate that robust evaluation is unfeasible, or suggest that a significantly different evaluation approach is required. The motivation for this approach is that to achieve a sufficient sample size, it is more cost-effective to include data from the pilot

[1] https://www.cebc4cw.org/program/the-oregon-model-parent-management-training-pmto-2/
[2] https://guidebook.eif.org.uk/programme/generation-pmto-group
[3] Forgatch, M. S., Patterson, G. R., Degarmo, D. S., & Beldavs, Z. G. (2009). Testing the Oregon delinquency model with 9-year follow-up of the Oregon Divorce Study. *Development and psychopathology*, *21*(2), 637-660.
[4] YEF's toolkit of existing research on parenting programmes highlights a lack of evidence about the direct impact of parenting programmes on crime and violence.

trial in the main, efficacy trial rather than recruit, deliver the programme to, and collect outcome data from an additional set of programme participants

To design this project, the Lab has collaborated with two partners. The first is ISII (Implementation Sciences International, Inc.), a research-based, non-profit organisation based in the USA, which implements the GenPMTO programme, in partnership with the programme developers. ISII also trains community practitioners in its use across the world. The second partner is Barnardo's, the UK's largest children's charity, and the delivery partner for the project. Barnardo's bring significant expertise and experience in delivering similar services and working with vulnerable families.

About the rationale for this evaluation

Whilst there is evidence of GenPMTO's impact on a range of outcomes, this is based on studies conducted outside the UK. There is a well-established phenomenon of programmes being transported into new countries and not demonstrating effectiveness when trialled in their new setting, and there are several examples in the UK of failed replication of programmes.[5,6,7] One often cited explanation for the high volume of null result trials in the UK, in the context of transporting programmes from overseas, is that these programmes haven't undergone the necessary formative work of ensuring they are appropriate, deliverable, and evaluable within the UK context.[8] These are necessary preparatory steps to a full-scale impact evaluation, and mitigate the risk of expending resources on an extensive trial before the programme has been adapted to the domestic context, which would ultimately yield uninformative results.

By undertaking a staged evaluation approach to GenPMTO, this enables gathering preliminary data on the programmes' feasibility of delivery and evaluability. This facilitates being able to refine the programme and approach to evaluation, to apply learnings from early stages to design a more robust full-scale efficacy trial.

Having established feasibility and acceptability in Stages 1 and 2, we will now investigate evaluability as part of Stage 3. If this is established, there is a strong case to proceed to a full-scale efficacy trial on GenPMTO in the UK, given the strength of its existing evidence, cultural similarity with the European countries where evidence has already been gathered in, and its potential impact on reducing arrests post-delivery.

While we will publish the findings of stages 1 and 2 separately, we identified that:

- The intervention was acceptable to both practitioners and caregivers, although some felt that the content was not always appropriate or optimally helpful for parents with older children.
- It was possible to recruit and retain caregivers in the intervention, although far larger numbers will be required for the randomised control trial, and there is more work to do to communicate the project well to referrers and support them.
- It is possible to collect outcome survey data from parents - incentives for completing the survey, using multiple methods of communication (particularly Whatsapp, in addition to email) and a flexible approach to data collection (i.e. permitting self-completion for parents who cannot attend scheduled survey sessions) are crucial.

---

[5] Robling, M., Bekkers, M-J., Bell, K., Butler, C. C., Cannings-John, R., Channon, S. et al. (2016). Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial. *Lancet, 387*, 146-155.

[6] Humayun, S., Herlitz, L. Chesnokov, M., Doolan, M., Landau, S. and Scott, S. (2017). Randomized controlled trial of Functional Family Therapy for offending and antisocial behavior in UK youth. *Journal of Child Psychology and Psychiatry 58*(9), 1023-1032.

[7] Fonagy, P., Butler, S., Cottrell, D., Scott, S., Pilling, S., Eisler, I. et al. (2018) Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): a pragmatic, randomised controlled, superiority trial. *The Lancet Psychiatry, 5*(2), 119-133.

[8] Lendrum, A., & Humphrey, N. (2012). The importance of studying the implementation of interventions in school settings. *Oxford Review of Education, 38*(5), 635-652.

These findings have fed into the design of Stage 3 and the approach we will take to delivery and evaluation, which will be further explored as part of this work.

About this protocol

This document is the evaluation protocol for Stage 3 - Pilot trial. Previously we published a protocol for Stages 1 and 2 - the Adaptation and Feasibility Study stages. Our rationale for publishing separate protocols is that the design of the latter stages of the evaluation has and will be influenced by what we learn in prior stages.

2. Intervention

Intervention overview

Background

GenPMTO is a targeted parenting programme for families with children and young people aged 3 -18, at risk of behaviour problems. It is based on the theoretical work of Gerald Patterson and colleagues at the Oregon Social Learning Center.[9] GenPMTO aims to improve school functioning, social relationships, and prevent involvement in criminal justice and substance use.

Based on robust evaluations, the programme has a strong track record of improving a range of important outcomes for parents and young people, such as reducing child externalising and internalising problems. The Early Intervention Foundation (EIF) awarded GenPMTO an evidence rating of 4.[10] Level 4 is the highest rating on the EIF strength of evidence scale, identifying programmes with evidence of a long-term impact and multiple rigorous evaluations. The programme has been found to positively impact both on child mental health and wellbeing (through improved child adjustment, social competence, emotional regulation, and a reduction in internalised behaviour problems), and on crime, violence and antisocial behaviour (through reduced police arrests, child conduct problems, externalising behaviour problems, antisocial-aggressive behaviour, delinquency and criminal behaviour, and improved social/prosocial behaviour). In terms of magnitude and longevity of impact, EIF calculated GenPMTO is associated with an 'improvement index' ranging from 11-22, with positive impacts identified as long as 8.5 years after programme completion.[11]

The programme has also demonstrated successful results when adapted to be delivered to culturally specific populations. For example: In the USA, the programme has been adapted for Latino/a families, and in Norway, the programme has been adapted for Somali and Pakistani mothers. Since the first international implementation in Norway in 1999, GenPMTO has been delivered to more than 50,000 families from a diversity of circumstances, socioeconomic backgrounds and cultures.

---

[9] Fisher, P. A., & Gilliam, K. S. (2012). Research into theory into practice: an overview of family based interventions for child antisocial behavior developed at the Oregon Social Learning Center. *Clin Salud, 23*(3), 247-259.

[10] https://guidebook.eif.org.uk/public/files/pdfs/programmes-generation-pmto-group.pdf

[11] An index of 22 means one would expect the median participant in the comparison group who did not receive the intervention (i.e., someone for whom 50% of their peers have better outcomes and 50% have worse outcomes), to improve to the point where they would have better outcomes than 72% and worse outcomes than 28% of their peers, if they had received the intervention.

Delivery of GenPMTO

GenPMTO can be delivered to caregivers on an individual family basis, or to groups of caregivers - however the focus of this project is the group-based version.[12,13] GenPMTO can also be delivered on-line, or in-person. More than one caregiver of a child or young person is able to be involved in the programme.

The programme delivery model is to train a first 'generation' of practitioners in a given context to full certification by ISII. A select set of these certified practitioners are trained to train subsequent cohorts of practitioners, if the programme is to be continued and/or scaled in that context.

Practitioners attend a number of workshops both before and during the delivery of the programme, with ISII staff supporting and providing feedback to practitioners throughout. ISII staff review practitioner delivery of the programme in order to grant full certification to practitioners who have met the required standards of delivery.

Practitioners are trained to deliver the programme to groups of 12-15 caregivers over either 10, 12, or 14 weekly sessions (depending on the delivery context). As part of this project, we are delivering the 14 session version. These sessions are approximately 90 -120 minutes in length and delivered by 2-3 practitioners.

During the feasibility study stage of this project, practitioners were trained in delivering the 14-week programme only. While we aimed to deliver sessions to groups of 12-15 caregivers during the feasibility study stage, due to recruitment and retention issues some groups ended the programme with as few as 5 or 6 caregivers.

Programme topics and skills

Delivered across 14 weekly sessions, the programme covers 14 essential topics, with sessions being agenda-driven, responsive, and focusing on skill building to promote effective parenting during times of transition. Practitioners use active teaching approaches (such as group problem-solving, role-play, and video modelling) to support caregivers in using positive parenting strategies at home. The topics on which caregivers receive training are as follows:

1. Working through change
2. Encouraging cooperation
3. Teaching positive behaviour
4. Observing emotions
5. Regulating emotions
6. Active communication
7. Setting limits
8. Following through

[12] We use the term 'caregivers' to refer to those individuals who are primarily and most consistently responsible for the care of a child. This is a broad term which includes parents, but also includes other individuals, such as grandparents, aunts, uncles, foster parents, etc. GenPMTO is designed to be delivered to a range of different types of caregivers and is not limited strictly to biological parents.

[13] Training of practitioners for group-based delivery of the programme can be more easily adapted for individual-based delivery than vice-versa, and so group-based delivery was agreed with delivery partners and programme as a starting point.

9. Communicating with children
10. Problem solving
11. Encouraging cooperation: incentive charts
12. Monitoring children's activities
13. Promoting school success
14. Putting it all together

The programme seeks to develop the following core skills with caregivers:

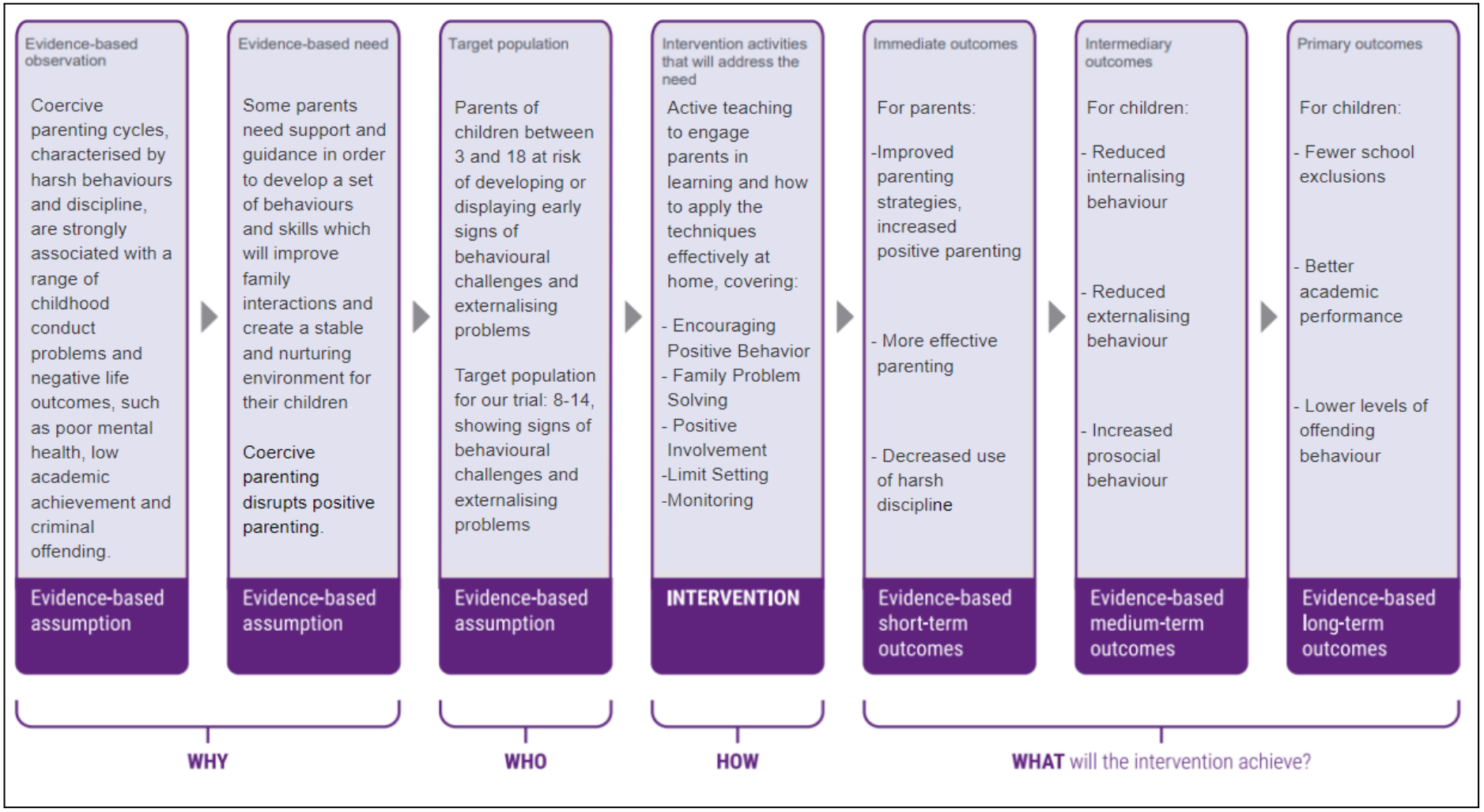
- Encouraging positive behaviour
- Limit setting with mild consequences
- Family problem-solving
- Positive involvement
- Supervision/monitoring
- Identifying/regulating emotions
- Promoting active communication
- Promoting success at school

An overview of GenPMTO using the TIDieR framework can be found in Annex A.

Figure 1 below details a high-level Theory of Change for the intervention.

Intervention theory of change

Figure 1 - GenPMTO theory of change



How the intervention compares to other services and business-as-usual

Parenting programmes such as GenPMTO focus on enhancing parenting practices and behaviour. These can involve: developing and practising positive discipline techniques, learning age-appropriate child development skills and milestones, promoting positive interaction between parents and children, and locating and accessing community services and supports.

Well-known examples of parenting programmes in the UK include Triple P, Incredible Years, and Strengthening Families. These programmes typically involve educational elements relating to child development, as well as training elements which support parents to develop specific skills. An initial survey of the three London boroughs (Brent, Tower Hamlets, and Barking & Dagenham) in which GenPMTO is anticipated to be delivered suggests that other such parenting programmes are currently being delivered in these areas.

To the best of our knowledge, while there is strong evidence that other parenting programmes can be effective at reducing behavioural difficulties (which are associated with later involvement in violence), there are limited examples of parenting programmes with robust evidence of reducing crime and violence itself.[14] GenPMTO is a rare example of a parenting programme that has measured these outcomes and robustly demonstrated improvements.

While some other interventions share these characteristics, GenPMTO can be distinguished from many other parenting programmes on the basis that a) it is more intensive, and is delivered over a longer period of time, b) has a stronger international evidence-base, and c) has evidence of reducing arrests. Overall, we are seeking to identify whether GenPMTO adds value above and beyond existing, business-as-usual provision in London.

[14] Gaffney, H., Farringdon, D.P. & White, H. (2021). Parenting Programmes: Toolkit technical report. Youth Endowment Fund: London. Retrieved from: https://youthendowmentfund.org.uk/wp-content/uploads/2021/06/Parenting-programmes-Technical-Report.pdf

3. Research objectives

This pilot trial primarily aims to assess the evaluability of GenPMTO in a UK context (by testing randomisation, data collection, gauging likely study attrition, etc.), and to determine whether the project could and should progress to a full-scale trial.

We also aim to collect high-quality outcome data - relating to parenting strategies and behaviours, and child behaviour - that could be analysed alongside outcome data collected in the subsequent (efficacy trial) phase to determine intervention impact. If the pilot trial is successful (based on the criteria discussed in Section 4), then there is a strong case to proceed to an efficacy trial and analyse this alongside the outcome data collected from the pilot trial to produce an estimate of the effect of the programme.

The pilot trial is designed to test the following questions:

- **Establishing evaluability** - Do we have enough confidence in the feasibility of a randomised controlled trial (RCT), particularly in terms of recruitment into evaluation, randomisation and outcome data collection, to justify extending the sample and progressing/continuing to efficacy trial?
- **Measuring outcomes** - Does GenPMTO show sufficient promise in terms of improvements in key outcomes to justify a subsequent efficacy trial?[15]

Alongside these we will continue to monitor:

- **Monitoring deliverability** - Can we retain participants randomised to GenPMTO in the intervention? Can we continue to deliver the programme with fidelity?
- **Monitoring acceptability** - Is the GenPMTO programme seen as acceptable and valuable by participants?

For more detail on research questions, please see Table 1 below.

[15] Note that the sample size in this pilot is unlikely to be big enough to make strong causal claims about effectiveness. However, it should be sufficient for us to produce descriptive statistics on key outcome variables and make an estimate of effectiveness with low levels of confidence.

Table 1 Research questions

Research objective	Focus	Research questions
Establishing **evaluability**	Recruitment (study)	*Referrals* ● How many families are referred to the project/study? ● What factors affect the volume of referrals? *Referral suitability* ● How many/what proportion of referred families are eligible? ● What are the most common reasons for families being deemed ineligible for the project/study? *Source of referrals* ● Which agencies and settings are referring families? ● How many families are being referred by each? Which agencies/settings contribute the most referrals and why? ● How many referrals are self-referrals (i.e. do not come through intermediaries, but in response to broader promotion/outreach approaches)? ● Are there differences between referring agencies/settings in terms of what proportion of participants are meeting inclusion criteria? *Recruitment and consent* ● How many eligible families can be successfully consented into the evaluation? ● How does this vary across referral source and participant characteristics, including: ○ Primary caregiver ethnicity ○ Primary caregiver gender ○ Primary caregiver age ○ Family SES ● Given the number of eligible families successfully consented into the evaluation over the time-period of the pilot, would a well-powered RCT be achievable at the efficacy stage over the currently planned time period, or would it need to be extended? ● What is the typical length of time between referral and randomisation?
	Randomisation	*Feasibility and adherence* ● Is the randomisation approach feasible (i.e. can we randomise straightforwardly as planned in this context, or are there unexpected barriers)? ● How many/what proportion of recruited participants complete baseline data surveys and are randomised? ● Is randomisation adhered to (i.e. is randomisation accidentally or intentionally subverted)? ● How could the approach to randomisation be adapted

		to increase feasibility and adherence? *Acceptability* ● Is randomisation acceptable to families? ● Is randomisation acceptable to practitioners? Do practitioners feel that evaluation activities (randomisation) impact the quality and ability to deliver the intervention well? In what way? ● Is randomisation acceptable to Boroughs and referring practitioners? ● What factors affect acceptability of randomisation? ● How could the approach to randomisation be adapted to increase acceptability?
	Control group services	● What alternative services or support (specifically parenting support) do the control group receive, if any? ● To what extent are alternative services similar to GenPMTO?
	Data collection & study retention	*Participant perceptions of data collection* ● How do participants feel about the questions asked in the outcome data survey? ● How do participants feel about the length of the outcome data survey? *Attrition rates and retention in evaluation* ● How many/what proportion of treatment group participants complete post-programme surveys? ● How many/what proportion of control group participants complete post-programme surveys? ● What data collection approaches work well in retaining treatment and control group participants? ● What factors affect attrition rates? Do attrition rates vary by participant characteristics? ○ Primary caregiver ethnicity ○ Primary caregiver gender ○ Primary caregiver age ○ Family SES ○ Baseline scores on parenting outcome measures ○ Baseline scores on child behaviour outcome measures. ○ Borough ○ Mode of delivery (online vs. in person).
	Effect sizes and sample size	● For each outcome, what is the point estimate of effect size, what is the confidence interval around it, and what implication would the range of plausible values have for the required sample size at the efficacy stage?

	Mechanisms and moderating factors	● Is it possible to collect the data that would be required to assess whether outcomes vary by: ○ Ethnicity (primary caregiver, or child, depending on focus of the analysis) ○ Gender (primary caregiver, or child, depending on focus of the analysis) ○ Age (primary caregiver, or child, depending on focus of the analysis) ○ Baseline outcomes (parenting outcomes, or child behavioural outcomes, depending on the focus of the analysis) ○ Family SES ○ Callous/unemotional traits ○ Attendance/engagement with the programme
Measuring **outcomes**	-	● For each outcome, what is the directional change, what is the point estimate of effect size, and what is the confidence interval around it? What does this suggest in terms of preliminary evidence that GenPMTO improves: ○ Child externalising behaviours ○ Child prosocial behaviour ○ Child internalising behaviour ○ Overall difficulties with child behaviour ○ Parenting strategies ○ Parental self-efficacy
Monitoring **deliverability**	Recruitment and take-up (programme)	● How many eligible families randomised to GenPMTO take-up the programme? ● Does take-up[16]/recruitment of families vary by: ○ Primary caregiver ethnicity ○ Primary caregiver gender ○ Primary caregiver age ○ Family SES ○ Borough ● Are there differences between referring agencies/settings in terms of what proportion of participants are taking part in the programme?
	Completion (programme)	● How many families attend each of the 14 GenPMTO sessions and how many complete GenPMTO? ● What factors affect programme completion? ● Does programme completion vary by: ○ Primary caregiver ethnicity ○ Primary caregiver gender ○ Primary caregiver age ○ Family SES ○ Baseline scores on parenting outcome

[16] We define take-up as attending at least one programme session.

		measures ○ Baseline scores on child behaviour outcome measures. ○ Borough ○ Mode of delivery (online vs. in person). ● Are there differences between referring agencies/settings in terms of what proportion of participants complete the programme?
	Fidelity	● Is the programme being delivered with fidelity to the programme design? ● What factors affect delivering the programme with fidelity? ● What are the key barriers and facilitators to delivering the programme well? ● What variations in delivery are appropriate for effective implementation? How does the originally planned balance between consistency and flexibility work in practice?
	Cost	● What is the average cost of delivering a GenPMTO group and the unit cost of delivering GenPMTO? ● To what extent does this vary across groups, and what drives any observed heterogeneity?
Monitoring **acceptability**	-	● Is the GenPMTO programme seen as acceptable and valuable by caregivers? ● What factors affect acceptability? ● Is there evidence that the adaptations taking place in Stages 1 & 2 of this study have improved acceptability?

4. Monitoring and success criteria

We will use monitoring criteria throughout Stages 3 for two purposes:

1. To monitor if the project is proceeding as expected, allowing for us to make adjustments or pause the work if needed.
2. To make recommendations to the Youth Endowment Fund as to whether progression to Stage 4 (efficacy trial) should be pursued at the end of the pilot study.

We will use RAG (Red, Amber, Green) ratings to rate the progress of target criteria, on a monthly basis. Criteria meeting red or amber cut-off scores will prompt the following changes to our approach:

- Criteria with Amber ratings will indicate reviewing or adjusting delivery.
- Criteria with Red ratings will indicate pausing delivery for a period of time to carefully assess what changes would be required to justify resuming delivery.

The quantitative monitoring criteria used to monitor evaluability-related objectives are described in Table 2 below. While the criteria below offer guidance for the progression of the evaluation on the basis of quantitative assessments, these will also be complemented by qualitative measures, such as ongoing practitioner feedback and interviews with practitioners and caregivers. Overall, we will not establish a deterministic rule about how many green-, amber-, and red-rated criteria would justify/prohibit progression, but instead will use the criteria to support a balanced judgement, weighting the importance of each criteria in the round.

The monitoring criteria specified below were developed iteratively in consultation with both Barnardo's and ISII, and help to ensure that the study does not progress from one phase to the next without sufficient evidence of success.

Table 2. Monitoring and success criteria for pilot study

Research objective	Criterion	Description	RAG scores
Evaluability	Eligible referral volume	Number of months (out of 7 recruitment months) where at least one Borough achieves a 'critical mass' of 30 eligible referrals where we can initiate the study onboarding and randomisation process for a cohort of parents/caregivers.[17] (*OR* a critical mass of 30 eligible referrals is achieved across all 3 Boroughs combined, provided those participants have indicated that they are content with online delivery).	Green: 5-6 months Amber: 3-4 months Red: 0-2 months
	Baseline data collection and randomisation	Number of participants who are randomised (and meet the requirements to be randomised, i.e. they're eligible, they've consented, they have complete referral information, and they complete baseline surveys).	Green: 92+ participants (over 2/3rds of 138 target) Amber: 47-92 (up to 2/3rds of target) Red: 0-46 (up to 1/3rd of 138 target)
	Evaluation retention - treatment	Proportion of participants randomised to the treatment group who submit the post-intervention survey. These thresholds are based on the EIF evidence standards guidelines.	Green: 90-100% of participants randomised to treatment Amber: 35-89% Red: <35%
	Evaluation retention - control	Proportion of participants randomised to the control group who submit the post-intervention survey.	Green: 90-100% of participants randomised to control Amber: 35-89% Red: <35%
Outcomes	Outcomes	Directional change in outcome variables for treatment and control.	Green: At least one outcome

[17] We are intending to conduct 7 tranches of randomisation and run 7 GenPMTO groups, one each month between December 2024 and June 2025. Based on conversion rates identified in the previous study stages, we anticipate that each tranche will require 30 eligible referrals (210 across all tranches), which we anticipate will result in approximately 20 participants completing baseline surveys and being randomised (138 across all tranches). We aim for no more than 15% attrition, with 17 completing post-test surveys (achieving our target of 120 completed pre- and post-test outcome surveys across the 7 cohorts). To run an in-person GenPMTO group, we will require 30 eligible referrals within one Borough. To run an online GenPMTO group, we will only require 30 eligible referrals combined across all 3 Boroughs (given that proximity to a physical location will not be an important factor for online delivery). However, these 30 participants need to indicate that their preference is to receive the programme online (or indicate that they have no preference).

		Where null results are those where 95% confidence intervals (CI) include 0, and positive/negative results are those where 95% CIs are entirely above or below 0 (depending on how the variable is coded and which direction is considered an improvement to the outcome).	measure indicates positive results, and none indicate negative results Amber: Null or mixed results Red: At least one outcome measure indicates negative results and no positive results
Deliverability	Programme completion	Proportion of participants randomised to GenPMTO who complete a sufficient proportion of the intervention (defined by developers as attending at least 8 sessions). We will monitor programme take-up and continued attendance throughout the study to identify if we are on track to meet these targets. Note that we will attempt to retain all randomised families in the evaluation, regardless of their attendance.	Green: 80-100% of participants randomised to GenPMTO Amber: 50-79% Red: <50%
	Fidelity	The proportion of practitioners achieving an aggregate mean Fidelity of Implementation Rating System (FIMP) score of at least 4, as rated by ISII programme experts. A FIMP score of 4 is considered to be the lowest score for what is considered acceptable implementation fidelity.	Green: 80-100% of practitioners Amber: 60-79% Red: <60%
Acceptability	Acceptability	Proportion of caregivers who broadly indicate that GenPMTO is acceptable and valuable.	Green: 80-100% of caregivers Amber: 60-79% Red: <60%

5. Design and methodology

Design

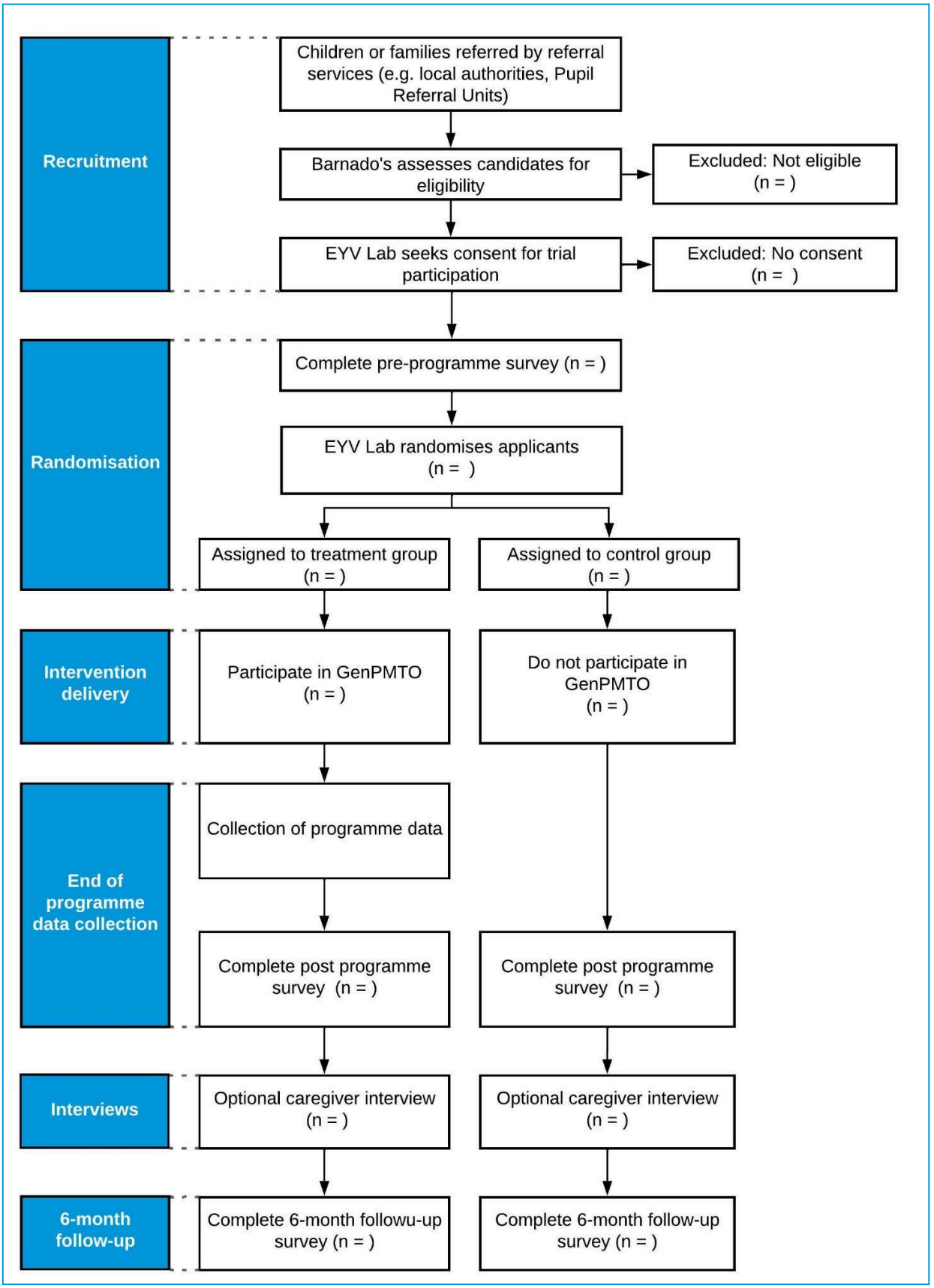
This pilot is designed as a two-armed superiority RCT, with the treatment arm receiving the GenPMTO programme, and the control group receiving business-as-usual services already in place within their Local Authority. The pilot study is designed to last approximately 9 months.

Process

1) **Referral** - As with the feasibility study, referral services (selected local authorities and settings/agencies within them) identify caregivers who may require support and refer them into the project (the 'parenting support project'). Information about the project is shared with families, and caregivers and their referrer will complete the referral form into the project, and consent is given to share this with Barnardo's and the Lab.
2) **Eligibility assessment** - Barnardo's will then determine whether the family in question are eligible for the project (and eligible to receive the GenPMTO intervention, should they be randomly assigned to it). If they are not eligible, they will not be included in the project and randomised, however they may be referred to pre-existing 'business as usual' local services.
3) **Consent and baseline data collection** - If a family is eligible for the project, the primary caregiver will be approached to consent to participation in the evaluation in the first instance by their referrer and talked through the study information sheet. When a sufficient number of eligible and consenting families are identified, the Lab will get in contact to attempt to collect baseline outcome data. For families that express a preference to receive their intervention in-person, they will be invited to an in-person group survey session where a Lab researcher will introduce the survey and support participants in completing it. For families that express a preference to receive their intervention online, these sessions will be conducted online over video conference software. Families that are not able to attend a scheduled survey session (either in-person, or online) will be permitted to self-complete their survey online at a time of their choosing, but will be able to text or call a Lab researcher for support. At this stage the research team will talk caregivers through the information sheet again, provide an opportunity to ask questions, and ask caregivers to reaffirm their consent to be involved. Those who do not consent to participation in the evaluation or do not provide baseline data will not be included in the project and randomised.
4) **Randomisation** - Families (one primary caregiver per family) completing baseline surveys will then be randomised by a researcher at the Lab. Barnardo's will be notified who has been randomised to the GenPMTO group and then reach out to them to schedule their GenPMTO programme. Boroughs will be notified who has been randomised to the control group - these participants will not receive GenPMTO but may be referred to pre-existing 'business as usual' local services.
5) **Post-test data collection** - We will attempt to collect outcome data from all families, regardless of whether they were randomly assigned to treatment or control groups, and regardless of how much of the programme they complete (an 'Intention-to-Treat' design). We will reach out to all families approximately 14-weeks after their baseline survey, after the GenPMTO intervention has concluded for participants randomised to the treatment group.
6) **Follow-up data collection** - As with other evaluations of GenPMTO, 6-months after the post-test survey, we will attempt to collect the same set of outcome data. We will encourage caregivers to self-complete these surveys in their own time.
7) **Data archiving** - At the beginning of a caregiver's involvement in the study, we will notify them

that we intend to submit their data to YEF's data archive, and seek their consent for this. Caregivers will need to consent to this to be eligible to participate in the study and be randomised. This will permit other researchers to follow up key outcomes (including offending) using administrative data years into the future, and identify GenPMTO's long-term impact.

Figure 2 - Participant flow diagram for pilot and efficacy trial


Recruitment
Children or families referred by referral services (e.g. local authorities, Pupil Referral Units)
Barnado's assesses candidates for eligibility
Excluded: Not eligible (n =)
EYV Lab seeks consent for trial participation
Excluded: No consent (n =)
Randomisation
Complete pre-programme survey (n =)
EYV Lab randomises applicants (n =)
Assigned to treatment group (n =)
Assigned to control group (n =)
Intervention delivery
Participate in GenPMTO (n =)
Do not participate in GenPMTO (n =)
End of programme data collection
Collection of programme data
Complete post programme survey (n =)
Complete post programme survey (n =)
Interviews
Optional caregiver interview (n =)
Optional caregiver interview (n =)
6-month follow-up
Complete 6-month followu-up survey (n =)
Complete 6-month follow-up survey (n =)

Randomisation and control group

Randomisation will take place at the level of the family (or, more specifically, parent-child dyads). For the purposes of data collection, we will collect data from one caregiver (the 'primary' caregiver who spends the most time with the child or young person) and one 'index child' (who meets the eligibility criteria and is the focus of the referral).

In cases where there is more than one primary caregiver, we will randomly select one caregiver to be the focus of data collection. In cases where there is more than one eligible child or young person nested within a single family, we will randomly select one child or young person to be the index child.
It is possible that two caregivers/parents from the same family may attend GenPMTO sessions, if their family is randomised to GenPMTO (although we will monitor this, and may limit this if we find it is limiting our ability to reach as many families as possible by occupying limited space available in the programme). When this occurs, the primary caregiver (or randomly selected primary caregiver) will remain the focus of data collection at post-test assessment. The same caregiver and the same child/young person will be the focus of data collection at baseline and post-test assessments.

We intend to randomise participants in 7 tranches of 30 eligible participants:

- For some tranches, randomisation will occur *within* Borough - When we successfully recruit 30 participants *within a borough* who indicate their preference is to receive the programme *in-person*, we will use simple randomisation to assign these participants to treatment or control.
- For other tranches, randomisation will occur *across* Borough - When we successfully recruit 30 participants across *all boroughs* who indicate their preference is to receive the programme *online*, we will use simple randomisation to assign these participants to treatment or control.
- Randomisation, then, is effectively stratified by Borough except for participants who wish to receive their service online. We have taken this approach because it will enable us to randomise and begin delivery for a subset of participants (those who indicate they are content with online delivery) faster and reduce waiting lists, as we will not need to wait for 30 eligible referrals within a *single* Borough to initiate these tranches.

Randomisation will take place after families are referred to the study, assessed for eligibility, and after baseline assessment. Caregivers will not be blind to treatment allocation. Randomisation will be implemented using a random number generator (using Stata 16.0).

This will not be a waitlist control study, as this pilot study is designed to test the evaluability of (and if progression to the next stage is achieved, feed into) long-term assessment of programme impact. Boroughs have been advised that control group participants can be supported with other services (i.e. business-as-usual), as they would have done in the absence of this project. This may range from no support, to light-touch support, to formalised parenting programmes like GenPMTO (depending on what services are available within a given Borough, the way the Borough standardly refers families to services, and the needs of the family).

Participants and sample size

Inclusion and exclusion criteria for families

We have carefully considered the inclusion and exclusion criteria for caregivers participating in the pilot study (and have tested these as part of the prior feasibility study). Our aims with these criteria are to:

1) Include caregivers who stand to benefit most from participation in the project and, potentially, in GenPMTO.
2) Exclude any caregivers for whom the project and/or programme may not be best suited, or may be at an increased risk of harm if they were to participate.

Inclusion criteria

Caregivers are eligible to participate in the study (and to receive GenPMTO, should they be randomised to it) if they:

- Have a CYP between the ages of 8-14.
- Are the primary caregiver (i.e. spend the most time with the CYP and are available to care for them).
- Live within one of the boroughs in which this project (the 'parenting support project' and evaluation of GenPMTO) is currently operating.

And, if one of or more of the following is present:

- CYPs have engaged in criminal behaviour, such as breaking the law or "offending behaviour" for both non-violent and violent crimes.
- CYPs have engaged in violent and challenging behaviour (including within the home, e.g. against parents and/or siblings).
- CYPs have been reported as bullying other individual(s) in or outside of school settings.
- CYPs have low attendance at school (<50% within the last academic year).
- CYPs have been excluded from school within the last academic year.
- CYPs are engaged in substance abuse/misuse (e.g. drugs, alcohol)
- CYPs are at risk of involvement by gangs.
- CYPs are at risk of exploitation, or negative influence, by criminal peers.
- CYPs have a sibling(s) that has entered into the criminal justice system.

Exclusion criteria

Caregivers will be excluded from the study if at least one of the following are present:

- Caregiver(s) have received a parenting programme in the last two months, or are currently receiving one.
- Caregiver(s) and CYP does not have working proficiency in English, such that participation in research activities (and GenPMTO, should they be randomised to receive it) would be unfeasible.
- Family has plans to move out of the borough within the 14 week delivery timeline, and thus may not be available for full delivery of GenPMTO should they be randomised to receive it.
- Severe developmental delay for caregiver or CYP which may prevent caregiver from attending GenPMTO delivery sessions and implementing GenPMTO parenting strategies (should they be randomised to receive it), or participating in evaluation.

- Caregiver(s) and/or CYPs are actively homicidal, suicidal or psychotic.[18]
- Problem sexual behaviour is the central behavioural concern for child/young person.[19]
- Significant child protection concern (i.e. basic needs of children are not being met by caregivers.[20]

Sample size

The pilot trial will seek to recruit approximately 100-120 families across both the treatment and control arms, across three London boroughs.

The primary aim of the pilot trial - on its own - is to test the feasibility of evaluating GenPMTO in the UK, on a relatively small scale. Subsequently it is unlikely that the pilot trial - on its own - will detect a statistically significant effect size. However, as previously noted, this pilot trial is intended to be an internal pilot trial, should the project progress to the next stage. This would involve extending delivery by a further 9-12 months, delivering the intervention to an additional cohort of families, and analysing their outcome data to make a robust assessment of the impact of GenPMTO. The new outcome data collected would be analysed alongside the outcome data collected from this pilot trial to provide an estimate of the effect of the programme. Therefore this pilot is intended to achieve *half* the sample size that would be required for a fully-powered efficacy trial. The following describes our rationale for the sample size of the *overall trial* (pilot stage + extended efficacy stage).

Overall, we note that previous evaluations of GenPMTO that have successfully identified effects on parent and child outcomes have involved samples of 96 to 238, and therefore this overall project would be similar in size or larger than previous successful efficacy studies. Given this, we are confident that this study will be sufficiently powered. In terms of **child outcomes:**

- Meta-analyses of the relevant academic literature on GenPMTO indicate an anticipated small effect size of approximately 0.148 (*Hedge's g*) on child outcomes.[21]
- However, there is a great deal of heterogeneity in the effect sizes observed for child outcomes in this meta-analysis, with the overall meta-analytic average including all child outcomes, across GenPMTO delivered in a multitude of ways, to all populations and in all circumstances.
- In general, studies that make use of parent-report measures, look into the group-format programme, and investigate samples at higher levels of risk/need report **larger effects overall** (across parent and child outcomes). All of these characteristics hold for our proposed design, therefore there is reasonable expectation that the expected effect size for this trial would be higher than the meta-analytic average.
- In the context of large heterogeneity of effect across the many studies evaluating GenPMTO, it is instructive to look at individual studies which most closely match the proposed design and circumstances of this proposed trial, which identify much larger effects:
    - Of the 3 evaluations included on EIF's Guidebook and assigned the highest ratings in terms of methodological robustness, one was conducted in Europe (Norway).[22] This study (Kjøbli et al., 2013) investigated a sample of children with a mean age of 8.56, and a higher-risk sample of children who scored above the cutoff for conduct problems on the Eyberg Intensity Scale.

[18] This exclusion is considered to indicate requiring escalation/referral to appropriate services (rather than enrollment in GenPMTO or a business-as-usual parenting service)
[19] As above.
[20] As above.
[21] Cai, Q., Chan, A. C., Lee, S. K., Marsalis, S., & Gewirtz, A. H. (2022). Effectiveness of GenerationPMTO to promote parenting and child adjustment: A meta-analytic review. *Clinical Child and Family Psychology Review*, 1-18.
[22] Kjøbli, J., Hukkelberg, S., & Ogden, T. (2013). A randomized trial of group parent training: Reducing child conduct problems in real-world settings. *Behaviour research and therapy*, *51*(3), 113-121.

    - In this study, effect sizes of **.42** were identified on the Eyberg Intensity scale at post-test. Another measure (the Merrell externalising scale) was also used to assess externalising problems, and found an effect size of **.15** at post-test and **.39** at follow-up.
- Using pre-post correlation statistics reported in one GenPMTO evaluation, we have been able to conduct power analyses that take into account the improvements in statistical power resulting from the inclusion of covariates (a pre-test covariate of conduct problems/externalising behaviour). Please see the results of our adjusted power analyses in the tables in Annex B.
- Our proposed trial of n=200-240 participants would be sufficiently well-powered to detect the majority of these benchmark effects from the comparison trial described in bold above. We are confident that the inclusion of additional covariates would reduce the minimal detectable effect size (MDES) of our planned study even further.
- We will examine data (effect sizes and confidence intervals) acquired from the stage 3 pilot trial and determine whether we may require a larger sample for the stage 4 efficacy trial.

In terms of **parent outcomes**:

- Meta-analyses of the relevant academic literature on GenPMTO indicate an anticipated effect size of approximately 0.24 (*Hedge's g*) on parenting outcomes.[23]
- However, as indicated above, there is a great deal of heterogeneity in the effect sizes observed for in this meta-analysis. In general studies that make use of parent-report measures, look into the group-format programme, and investigate samples at higher levels of risk/need report larger effects. All of these characteristics hold for this proposed design, therefore there is reasonable expectation that the expected effect size for this trial would be higher than the overall meta-analytic average. For instance:
    - The meta-analytic average for parent outcome effects based on self-report measures is **0.367** - requiring a sample size of 236 for acceptable levels of power.
    - The meta-analytic average for overall parenting outcomes is **0.423** - requiring a sample size of 178 for acceptable levels of power.
- In the context of large heterogeneity of effect across the many studies evaluation GenPMTO, it is also instructive to look at individual studies which most closely match the proposed design and circumstances of this proposed trial, which identify much larger effects:
    - Of the 3 evaluations described above, one was conducted in Europe (Norway).[24] This study (Kjøbli et al., 2013) investigated a sample of children with a mean age of 8.56, and a higher-risk sample of children who scored above the cutoff for conduct problems on the Eyberg Intensity Scale.
    - In this study, effect sizes of **0.88** and **0.87** were identified on positive parenting and harsh discipline respectively, which we will be measuring in this trial.
- Unfortunately we have not been able to conduct power analyses that take into account the improvements in statistical power resulting from the inclusion of covariates, as the necessary information is not reported in previous evaluations of GenPMTO. However, please see the results of our unadjusted (and therefore conservative, i.e. over-estimating the required sample) power analyses in Annex B.
- Our proposed trial of n=200-240 participants would largely be sufficiently well-powered to detect these benchmark effects from the meta-analysis and from the comparison trial described in bold above (even without taking covariates into account). We are confident that

[23] Cai, Q., Chan, A. C., Lee, S. K., Marsalis, S., & Gewirtz, A. H. (2022). Effectiveness of GenerationPMTO to promote parenting and child adjustment: A meta-analytic review. *Clinical Child and Family Psychology Review*, 1-18.
[24] Kjøbli, J., Hukkelberg, S., & Ogden, T. (2013). A randomized trial of group parent training: Reducing child conduct problems in real-world settings. *Behaviour research and therapy*, *51*(3), 113-121.

the inclusion of additional covariates would reduce the minimal detectable effect size (MDES) of our planned study even further.

- As above, please note that we intend to use data acquired from the stage 3 pilot trial to inform a new set of power analyses before the stage 4 efficacy trial proceeds. This will provide us with a more accurate estimate of the required sample size, and will allow us to alter our recruitment targets accordingly in the fourth stage, to ensure that we recruit sufficient participants.

Outcome measures

Type of outcome	Outcome measured	Instrument	Completed by	Number of items	Age suitability (young person)	Subscales to be used	Scoring	References
Primary	Externalising Behaviour	Eyberg Child Behaviour Inventory (ECBI): Intensity Score	Parent report of child	36	Parents of children and young people between 2-16 years old.	All subscales. While the **intensity score** is our core primary outcome, we will also report against the **problem score**.	**Intensity score:** A score ranging from 36-252, indicating how often a series of challenging behaviours occur. Items are assessed on a 7-point Likert scale, from 1 ('Never Occurs') to 7 ('Always Occurs'). Higher scores indicate higher frequencies of challenging behaviours. **Problem score:** A score ranging from 0-36, indicating whether or not a parent considers a series of challenging behaviours as a problem. Items are assessed ('Yes' = 1 , 'No' = 0). Higher scores indicate that more challenging behaviours are a problem.	Eyberg, S., & Ross, A.W. (1978).[25]
Secondary	Internalising behaviour	Strengths and	Parent report of	25	Parents of children	Some subscales	**Internalising score:** A score ranging from 0-20, generated	Goodman,1997[26]

[25] Eyberg, S., & Ross, A.W. (1978). Assessment of child behavior problems: The validation of a new inventory. Journal of Clinical Child Psychology, 7, 113–116.
[26] Goodman R (1997) The Strengths and Difficulties Questionnaire: A Research Note. Journal of Child Psychology and Psychiatry, 38, 581-586.

		Difficulties Questionnaire (SDQ)	child		and young people between 4-17 years old.	including: ● Emotional symptoms ● Peer relationships problems	by summing scores from the emotional and peer problems subscales. Most items are assessed on a 3-point Likert scale ('Not true' = 0, 'Somewhat true' = 1, 'Certainly true' = 2), although some are reverse coded. Higher scores indicate more challenging behaviours.	
Secondary	Parenting strategies and use of harsh discipline	Parenting Practices Interview (PPI) 2019 version	Parent self-report	64	Parents of children and young people between 3 and 12 years old.	All subscales (positive and negative parenting/discipline).	**Total scores** range from 64 to 448. Items are assessed on a variety of different scales, although most are assessed on a 1-7 Likert Scale. Overall, higher scores indicate improved parenting practices	Parenting Practices Interview (PPI) 2019 version
Secondary	Parental self-efficacy	Parental Locus of Control - Short Form Revised (PLOC-SFR)	Parent self-report	24	Parents of children and young people between 5 and 11 years old.	All subscales.	**Total scores** range from 24 to 120. Items are assessed on a 1-5 Likert scale. Overall, higher scores indicate greater parental self-efficacy.	Hassall et al., 2005 [27]

In addition, we will be collecting the 24-item parent-reported version of the Inventory of Callous-Unemotional Traits (ICU-Parent). This will inform separate, exploratory analysis of variables that may moderate intervention impact, and will not inform the main findings of the pilot study or the subsequent efficacy trial.

[27] Hassall R.; Rose, J. and McDonald, J. (2005). Parenting stress in mothers of children with an intellectual disability: the effects of parental cognitions in relation to child characteristics and family support. J Intellect Disabil Res.

Methods and data collection

We will be conducting quantitative research activities with the following key data sources:

1) **Caregiver surveys:** surveys will be conducted with caregivers at baseline, at post-intervention, and at 6-month follow-up points (i.e. 6 months after programme completion) with both intervention and control groups. These measurement timings have been used in previous robust trials of Generation PMTO. Based on the prior feasibility study, we expect caregivers to take 30-45 minutes to complete the survey.
2) **Administrative data:** including referral data, consent form data, and programme attendance sheets.
3) **Programme fidelity data:** assessments of fidelity provided by the programme developers (using the Fidelity of Implementation Rating System).

Quantitative and qualitative methods and data collection summary (orange indicates quantitative methods, blue indicates qualitative)

Purpose	Focus	Source	Data collected	Data analysis
Establishing **evaluability**	Referral into study	Administrative data (referral data)	Number of referrals	Descriptive statistics
	Referral suitability	Administrative data (referral data)	Number of referrals meeting eligibility criteria.	Descriptive statistics
	Source of referrals	Administrative data (referral data)	Number of referrals from each source.	Descriptive statistics
	Promotion/outreach approaches	Administrative data (referral data)	Number of referrals arising from each outreach approach.	Descriptive statistics
	Recruitment into study	Administrative data (consent form data) and caregiver survey	Number of participants consenting into the study and providing baseline data.	Descriptive statistics
	Retention in study	Caregiver survey	Proportion of randomised participants completing post-test and 6-month follow-up surveys	Descriptive statistics
	Compliance with	Administrative	Number of	Descriptive

	randomisation	data (programme attendance sheets)	instances a participant not randomised to GenPMTO attends a session.	statistics
	Acceptability of randomisation	Caregiver survey	Likert-scale and open text responses to questions about randomisation.	Descriptive statistics
		Caregiver interviews	Caregiver's perceptions of the acceptability of randomisation.	Thematic analysis
		Practitioner interviews	Practitioner's perceptions of the acceptability of randomisation.	Thematic analysis
		Referrer interviews	Referrer's perceptions of the acceptability of randomisation.	Thematic analysis
	Control group services	Caregiver survey	Programmes control participants indicate they have attended (multiple choice question listing local services + a free-text 'other services' field).	Descriptive statistics
	Acceptability of data collection	Caregiver survey	Likert-scale and open text responses to questions about data collection.	Descriptive statistics
		Caregiver interviews	Caregiver's perceptions of the acceptability of data collection.	Thematic analysis
Measuring **outcomes**	Outcome data at pre-test, post-test and 6-month follow-up	Caregiver survey	A variety of surveys/questionnaires, specified in sections above.	Pre-specified statistical analysis

	Demographic variables for covariates and to inform subgroup analysis	Administrative data (referral data)	As specified in sections above.	Pre-specified statistical analysis
	Other variables to inform subgroup analysis and other exploratory analyses	Caregiver survey	As specified in sections above.	-
Monitoring **deliverability**	Programme take-up	Administrative data (programme attendance sheets)	Proportion of participants randomised to GenPMTO who attend a GenPMTO session.	Descriptive statistics
	Programme completion.	Administrative data (programme attendance sheets)	Proportion of participants randomised to GenPMTO who complete programme (complete at least 8 sessions).	Descriptive statistics
	Fidelity	ISII assessment	Programme developers assess sessions using the Fidelity of Implementation Rating System.	Descriptive statistics
Monitoring **acceptability**	Acceptability of programme	Caregiver survey	Likert-scale and open text responses to questions about programme delivery.	Descriptive statistics
		Caregiver interviews	Caregiver's perceptions of the acceptability programme delivery.	Thematic analysis
	Acceptability of programme adaptations	Caregiver survey	Likert-scale and open text responses to questions about programme adaptations.	Descriptive statistics

		Caregiver interviews	Caregiver's perceptions of the acceptability of programme adaptations.	Thematic analysis

Data analysis

The following section provides a high-level overview of our data analysis plan - we will provide a more detailed overview in a full statistical analysis protocol which will be published at a later date.

Feasibility, acceptability and evaluability analysis

The primary objective of the pilot trial is to establish the evaluability of GenPMTO, and to monitor deliverability and acceptability. This will be done based on descriptive analysis of quantitative data, as described in Table 2, which will be assessed based on the monitoring criteria discussed in Section 4 of this protocol. Analysis will be pre-specified in the full statistical analysis plan.

Effectiveness analysis

In addition to the analysis of the evaluability, deliverability and acceptability of GenPMTO, we will conduct analysis on the outcomes of the pilot study using the following approach.

This pilot trial - on its own - is not primarily designed to estimate effect sizes or evaluate the impact of the intervention in depth. However, we will conduct effectiveness analyses of the outcome data we acquire during the pilot to lay the groundwork for future evaluation and to inform our monitoring criteria and recommendations to YEF. These analyses may help us to understand the potential impact of the GenPMTO programme in supporting children at risk of youth violence in a UK context. However, we note that because this is a pilot study with a small sample size, we will have to interpret any statistical results with caution, and the pilot study on its own is unlikely to be sufficiently powered to detect an effect.

Based on YEF's outcomes framework, we have organised outcomes into primary and secondary groups, based on which outcomes are most predictive of youth violence and crime.[28] The primary outcome is:

- Reductions in child externalising behaviour (Eyberg Child Behaviour Inventory, assessed at post-test)

The secondary outcomes are:

- Increases in child prosocial behaviour (Strengths and Difficulties Questionnaire: prosocial subscale, assessed at post-test)
- Reductions in child internalising behaviour (Strengths and Difficulties Questionnaire: internalising score, assessed at post-test)
- Reductions in overall difficulties with child behaviour (Strengths and Difficulties Questionnaire: total difficulties score, assessed at post-test)
- Improvements in parenting strategies (Parenting Practices Interview, assessed at post-test)
- Improvements in parental self-efficacy (Parental Locus of Control - Short Form Revised, assessed at post-test).

[28] https://youthendowmentfund.org.uk/wp-content/uploads/2022/08/YEF-Outcomes-Framework-August-2022.pdf

- All outcomes assessed at 6-month follow-up.

All outcome data will be analysed using an intention to treat (ITT) analysis and linear (or logistic where relevant) regressions. We will collect pre-intervention outcomes for all families to increase power and adjust for regression to the mean. Our control vector will include ethnicity, gender, age, family SES, Borough and baseline outcomes (primary caregiver, or child, depending on focus of the analysis).

$$Y_i = \alpha + \beta_1 * Treatment + \beta_2 * Control\ vector + \varepsilon_i$$

In addition to the primary analysis of short- and medium-term outcomes, we will also:

- Conduct sub-group analyses to analyse differential evaluation recruitment and retention, and programme take-up and completion. The characteristics we examine will be pre-specified in the full statistical analysis plan.
- Enable long-term follow-up of offending data. We will collect the necessary data and submit this for archiving and linking according to YEF processes. Given the age of the children in this sample, and the appropriate time to analyse offending data likely being 4+ years after programme completion, this follow-up will be conducted outside of this project. This follow-up analysis may be conducted by the Lab, or by another organisation such as YEF, or commissioned by YEF.

Risks to study validity

Note that while the following risks are not problematic for the pilot per se (in terms of successfully assessing evaluability), they are risks for the overall impact evaluation if the project proceeds to the stage 4 efficacy trial.

Achieving insufficient statistical power

Although not a risk to internal validity per se, there is a risk that we do not acquire sufficient referrals to achieve half the sample size required for a fully powered trial, during this pilot study. This would reduce our ability to detect the impact of GenPMTO overall (once stage 4 is completed), even if there was an effect (i.e. a false negative).

We are aiming to acquire referrals and randomise in 6 (or more) tranches for the pilot study, triggering the onboarding and start-up process when we acquire 30 eligible referrals each month. If we do not reach 30 eligible referrals in each month, we will either need to: i) make the decision to initiate a group in that month with fewer participants than planned - and to compensate for this by committing to run additional cohorts as part of the pilot, or ii) delay the schedule to allow more time for referrals.

As well as limited referrals, study attrition may also impact our ability to achieve sufficient statistical power. This is discussed below.

Attrition (control and treatment group drop-out)

Attrition from the study may reduce the statistical power of our study and introduce bias. Some degree of attrition is to be expected. We will attempt to minimise the impact of this by:

- Providing incentives to the control and treatment groups (£20 per completed survey);
- Ensuring that evaluation activities are designed to be low-impact in terms of burden and time (including offering flexibility in terms of how surveys are completed);

- Where needed, utilising the relationships that Barnardo's and referring practitioners have built with participants to facilitate access and cooperation.
- We will inform participants with sufficient notice about planned survey activities.
- We will monitor rates of attrition, and whether there are differences in rates of attrition based on parent/caregiver ethnicity, gender, age, and in terms of baseline outcomes.

Contamination/spillover

It is possible that within the same geographic area (London borough), some spillover may occur, with caregivers receiving the treatment discussing parenting strategies and what they have learnt in GenPMTO with caregivers in the control group. The likelihood of spillover is deemed to be low however, given that the programme is covering a large geographical area with over 200,000 residents. Moreover, the programme is intensive and involves a variety of roleplay exercises and activities to support parents in developing and implementing new parenting strategies. We believe that any effects are likely to emerge as a result of participation in the full programme, and don't expect that sharing information from the programme would provide the same benefits. Finally, if spillover does occur, we believe this would lead to an improvement in outcomes for the control group, which would likely cause an underestimation of the effect size of the intervention, which is relatively tolerable as a risk (when compared to an overestimation).

Unrepresentative business-as-usual offer

We intend this pilot study to contribute outcome data that will permit, along with the stage 4 efficacy study, to an estimate of the impact of GenPMTO in the UK. Our overall aim is to determine whether GenPMTO is improving families' lives and adding value beyond the services parents/caregivers already access in London. There is a risk that, in the context of a study like this, that Boroughs overcompensate by offering control group participants an artificially boosted offer, by providing services similar to GenPMTO (i.e. similar aims, methods, and level of intensity of delivery). This would limit our ability to answer our key research question, and would reduce the likelihood of identifying an effect for GenPMTO.

Given variation in the level of need of families referred into the project, and variation in how, why, and by whom they've been referred, we would expect the business-as-usual for these families to naturally involve a mix of services, and for the business-as-usual offer not to be dominated by high-intensity programmes similar to GenPMTO. We have emphasised in all communications with Boroughs that we do not want control groups to receive an artificially restricted or artificially boosted offer in the context of this trial, and that control group participants should receive services that the Borough would have provided to these caregivers in the absence of the GenPMTO project. We will monitor what services the control group receive.

Qualitative research

As part of the feasibility study we conducted substantive qualitative research with caregivers, gauging their perception of the quality of the content and programme delivery, barriers to engagement and areas for improvement. For caregiver interviews in particular, we sampled caregivers to achieve a diversity of viewpoints (in terms of age, gender, ethnicity, educational attainment and other characteristics).

We also interviewed all GenPMTO practitioners and managers from Barnardo's, covering similar ground but also exploring their perceptions of the fidelity monitoring process, programme training and ongoing coaching. We also interviewed professionals referring families into GenPMTO, to explore barriers and facilitators to referring into GenPMTO and their perceptions of the programme.

Having learnt a great deal through these activities (and having found we reached data saturation with a relatively small number of interviews), we will conduct lighter-touch qualitative research as part of the pilot to focus on new research objectives associated with the pilot study (i.e. randomisation and evaluability), and to focus on new elements of the programme (i.e. adaptations to the programme made on the basis of the learnings of the prior feasibility study).

We will conduct:

- 2 caregiver interviews within each Borough (6 total), where caregivers received GenPMTO. These will be caregivers who have received GenPMTO. Interviews will be individual and we will aim for them to last one hour. We will aim to achieve a balance in terms of ethnicity and in terms of the age of their children.

Characteristic	**Quota**
Child age	3 families with children between 8 and 11 3 families with children between 12 and 14
Ethnicity	2 families where child is identified as White 2 families where child is identified as Black 2 families where child is identified as Asian

- 1 caregiver interview within each Borough (3 total), where caregivers were not randomised to receive GenPMTO. We will take a convenience approach to sampling.
- 2 practitioner interviews within each Borough (6 total). Interviews will be individual and we will aim for them to last 45 minutes. We will take a convenience approach to sampling.
- 2 referrer interviews within each Borough (6 total). Interviews will be individual and we will aim for them to last 30-45 minutes. We will aim to achieve a diversity in the participants we interview in terms of roles/teams.

Caregiver interviews (treatment group)

These interviews will explore:

- Perceptions of how the programme has impacted on them and their parenting.
- Barriers or facilitators to participating in the intervention.

- Perceptions of the age-appropriateness of GenPMTO content, to gauge the success of adaptations made in the feasibility study.
- Perceptions on the evaluation, particularly the process of being randomised, whether this was clear to them, and if they think any aspect of this could be improved.

Caregiver interviews (control group)

These interviews will explore:

- Perceptions on the evaluation, particularly the process of being randomised, whether this was clear to them, and if they think any aspect of this could be improved.

Practitioner interviews

These interviews will explore:

- Perceptions of how the programme has impacted on caregivers and their parenting.
- Barriers or facilitators to delivering the intervention.
- Perceptions of the age-appropriateness of GenPMTO content, to gauge the success of adaptations made in the feasibility study.
- Perceptions on the evaluation (particularly the process of randomisation preceding programme delivery), and whether they feel that the evaluation activities impacted subsequent delivery in any way, whether they picked up on any attitudes or beliefs relating to the evaluation in their work with parents/caregivers, and if they think any aspect of the evaluation could be improved.

Referrer interviews

These interviews will explore:

- Factors involved in making a referral decision.
- Ease of referral, including speaking to caregivers and referrer's role in supporting engagement.
- Their perceptions on making a referral into an RCT project specifically, whether the intentions of the work and the process of randomisation were clear, and whether they picked up on any attitudes or beliefs relating to the evaluation in their work with parents/caregivers.

Barnardo's manager interviews

These interviews will explore:

- Barriers or facilitators to managing the delivery of the intervention. Perceptions of what well and less well.
- Perceptions of the age-appropriateness of GenPMTO content, to gauge the success of adaptations made in the feasibility study.
- Perceptions on the evaluation (particularly the process of randomisation preceding programme delivery), and whether they feel that the evaluation activities impacted subsequent delivery in any way, whether they picked up on any attitudes or beliefs relating to the evaluation in their work with other stakeholders, and if they think any aspect of the evaluation could be improved.

To analyse the depth interviews, we will employ a version of the framework approach which is widely used in applied social research and draws on the approach set out by Ritchie et al (2014). This approach is similar to other widely used thematic analysis approaches and aims to derive meaningful themes and patterns from the qualitative data. However, rather than focusing on coding the data, this approach involves summarising, or 'charting' the data into a thematic framework. The strength of the

framework approach is that it enables systematic and comprehensive analysis of the complete data set in a manageable way. It involves the following steps:

1) *Transcription* - All interview recordings will be transcribed verbatim to ensure accuracy and facilitate subsequent analysis. Transcripts will be anonymised by assigning unique identifiers to each participant, replacing their names or any identifying information.
2) *Familiarisation with the data* - The research team will thoroughly read and familiarise themselves with key interview transcripts and observation notes - what is known as deep hanging out in the data. This step ensures that whatever headings are selected for the thematic framework are grounded in the data.
3) *Data management: Establishing initial thematic framework* - The first stage of data management will be for the research team to convene and discuss the possible themes that are emerging from the data, under which the data will be sorted. These themes will be both deductive (guided by the research questions and topic guides) and inductive (those that emerge from the data). Once the research team has agreed the key themes and sub-themes, these will be used to set up an initial thematic framework. For ease, this will be done in Excel/Google docs. The framework will be set up so that each individual sheet represents a theme and the columns within it represent the sub-themes. The rows represent individual participants. Each participant group will have their own thematic framework, so one for caregivers, another for practitioners etc. In some cases, the framework approach requires indexing and sorting of the data where the themes in the framework are used to annotate and label the data in the transcripts. However, since there will be a clear structure to the depth interviews, it is anticipated that the data will already be well ordered and this step will therefore not be needed.
4) *Data management: Charting* - Once the frameworks are set up, the data from each transcript will be 'charted' or summarised into them. The summaries will be written in the third person and aim to capture the key views of the participant under each of the themes represented by columns. The researchers doing the charting will remain as close as possible to the language used by the research participants. As the data is charted, researchers will identify key verbatim quotes from the transcripts and add these to the framework in italics to be used in the report if needed. Charting will be done by several researchers who will all read and quality assure each other's charting to ensure a consistent and comprehensive approach. Throughout the data management stage, researchers will be mindful of revisiting the thematic framework and adjusting it where needed. For example, adding new sub-themes that were previously not discussed or collapsing themes together where necessary.
5) *Analysis and interpretation* - The first stage of analysis is descriptive. This will involve looking at each theme in turn and exploring the range of views held under that theme with a view to developing categories. This will be done by grouping the views into clusters and exploring the properties of each of these clusters until clear categories can be developed. Given the nature of the feasibility study and the size and likely diversity of the sample, it is highly likely that the majority of the analysis will be descriptive and aim to clearly map out the range and diversity of views that exist within each participant population on the key areas relevant to the research questions. However, where possible the researchers will proceed to a higher level of analysis and aim to look for patterns and linkages in the data. This stage will be facilitated by the framework approach as it easily allows the researcher to look both within and across cases to see how different parts of the data set are connected. The sorts of patterns and linkages that might be explored include links between particular experiences of the intervention and how those link to views or outcomes, or the research team may explore links between particular characteristics of the participants and their views and experiences. Where possible, the

research team will then go on to look for explanations for the categories and linkages that have been found. Throughout the analysis process, the research team will remain in contact with each other, sharing and testing emerging findings and ensuring that the analysis process remains rooted in the data.

6. Cost evaluation - data collection and reporting

We will report the cost of delivering the intervention in the final report, following YEF costing guidance. We will:

- Use a bottom-up costing approach and break costs down into: prerequisites, set-up costs, and recurring costs.
- We will report the total cost for a typical single cohort receiving the intervention for one round of delivery and the costs per participant for one round of delivery, assuming full compliance. Depending on heterogeneity in costs across cohorts, we will either report average costs, or select a case which we think is most representative of the costs we expect to be incurred in future, typical rounds of delivery.

The organisations and practitioners involved in delivery are Barnardo's and the programme developer ISII. To report cost at the end of this study, we have produced a template for these organisations to complete, covering staff costs, equipment/materials costs, programme procurement costs, and buildings and facilities costs..

We expect most costs to fall within the following two categories:

- Staff cost: cost of practitioners, supervisors, and managers involved in delivering Barnardo's.
- Programme procurement costs: the cost for ISII to train and supervise staff, as well as to produce and provide the programme-specific materials required to deliver the programme.

7. Planned outputs

The outputs of this overall package of work and their timings are dependent on the outcomes of stage of the work. All findings will be published at the point the project concludes.

Regardless of the outcome of the pilot study, the Lab will provide a short report to the YEF outlining the key findings and observations of the pilot study. This will make a formal conclusion as to whether GenPMTO can be feasibly evaluated in the UK, and recommend whether the evaluation should progress to Stage 4 (efficacy stage).

If the pilot study concludes that GenPMTO is sufficiently feasible to evaluate in the UK and ready to progress to a full-scale RCT, we will not publish these findings externally at the completion of the pilot study. We will instead report the full findings of the pilot study at the end of the *entire* study, once the efficacy stage is completed. This report will contain findings and/or recommendations from the feasibility study, pilot trial and/or full-scale efficacy trial of GenPMTO. The Lab will share the findings more broadly where appropriate (and with the express permission from Barnardo's and YEF), such as

in presentations, blogs and the Behavioural Insight Team's (BIT) annual report. The primary audience of YEF's website is practitioners and researchers interested in reducing young people's involvement in Violence.

If the pilot study concludes that GenPMTO is not sufficiently feasible to evaluate in the UK and not ready to progress to a full-scale trial, then these findings will be published externally at the completion of the pilot study in the form of a publicly accessible report, which will include the key findings of the feasibility and pilot studies.

8. Ethics and data protection

Ethics

Overview

This trial was self-assessed as being high risk due to the inclusion of high-risk participants in the form of vulnerable young people. As a result we sought ethical approval from an independent panel of external experts with experience of working with vulnerable children and experience with safeguarding and child protection. We sought input from our panel twice - once, at the beginning of the feasibility study, and again during the transition from feasibility to pilot RCT.

The independent ethics review committee (ERC) reviewed the following information:

- Ethical review form
- Consent forms and information sheets for young people and parents/caregivers of young people
- Topic guides
- Outcome surveys
- Safeguarding and distress protocol
- A specification of the approach to randomisation and nature of intended control group.

The ERC discussed any issues raised by the research with The Lab with the aim of finding solutions that meet ethical requirements. The reviewers and the project manager agreed solutions to any outstanding issues, and the resulting changes to the way the project is being implemented have been included on the ethics form. The ERC was happy to approve the project with the inclusion of these amendments.

If there are substantial changes while the research or evaluation is being implemented, the ethics form will be revised and the revisions agreed with the ERC.

Informed Consent

All participants will be asked to provide written consent to participate in the pilot study for ethical purposes, before data collection, randomisation and interviews take place. Participants will be provided with an information sheet to inform them of what to expect from their involvement in the pilot study.

We will invite all caregivers to complete outcome surveys. All participants will be informed of their right to withdraw their consent at any point during interviews and/or data collection sessions. We will make

it clear to participants that we will use their information to inform the findings of our evaluation, which will be incorporated into a report, or other publicly publishable materials. However, no identifying information will be disclosed in any such materials.

We will also inform participants that they may be able to withdraw their data from the study, up until approximately 6 months after the end of the study. At this point their data will be deleted from our systems, and anonymised information will have already been incorporated into reporting, or other publicly publishable materials.

Safeguarding

Safeguarding means protecting the health, wellbeing and human rights of children and at-risk adults, enabling them to live safely, free from abuse, violence and neglect. During the pilot study, we will protect adults-at-risk by following a strict safeguarding and distress protocol. Before any interviewing or surveying with vulnerable groups, any researcher will:

- Undergo, and obtain, an **enhanced DBS check**.
- Complete the **NSPCC's Introduction to safeguarding and child protection training**.
- Review the **Nesta Group Safeguarding Policy** and the **GenPMTO Child safeguarding issues and Risk Assessment**.
- Review the **GenPMTO Safeguarding and distress protocol**.

If, during any research activity, a participant discloses anything that leads a researcher to believe that they themselves, or someone else, might be at risk of harm**,** they will follow these steps:

- **Step 1:** ***Is there an immediate risk of harm to the interviewee or others?***
  - If yes: they will call the police or other emergency services as soon as possible and follow up with an emergency report to the Nesta Group Chief People Officer (The Lab's Designated Safeguarding Lead) and the project's qualitative lead.
  - If no: proceed to step 2.
- **Step 2:** ***Establish an understanding of what has happened***
  - They will keep questions to the minimum necessary to ensure a clear and accurate understanding of what has been disclosed. The researcher will only ask questions to help establish whether the participant is at risk of harm. They will not make allegations or lead them to make allegations.
  - The researcher will ask the young person or at-risk adult whether anyone is aware of what they have disclosed e.g. are the parents/caregivers aware of it. If the concern is about a pre-existing mental health condition that is known to the caregivers or medical professionals (e.g. the adults GP), for example, this would not represent a safeguarding concern that would need to be reported.
  - *If you notice something concerning, which hasn't been disclosed:* they will ask open questions to establish if there is an explanation e.g. "that looks like a big bruise. Can you tell me what happened?"
- **Step 3:** ***Make a written record***
  - The researcher will note down what has been said, any physical evidence that is available including injuries or the personal state of the participant.
- **Step 4:** ***Inform the participant***
  - If the researcher considers that there is a risk, they will inform the participant that they need to tell the Lab's designated safeguarding lead. They will explain that a safeguarding lead is the person in an organisation that's responsible for dealing with concerns to people's safety.
- **Step 5:** ***Report the concerns***

- Nesta Group Chief People Officer (the Lab's Designated Safeguarding Lead) and the project's qualitative lead.
- The researcher will be available to the designated safeguarding leads to assist with further assessments, including whether cases need to be escalated to other parties such as the child protection services.

Data protection

We have followed appropriate data protection processes in accordance with BIT processes, including completing a Data Protection and Security Checklist and Data Protection Impact Assessment, which have both been reviewed and approved by BIT's legal team.

The Lab will store and handle all data securely and confidentially in line with requirements of the UK GDPR, and Data Protection Act (2018), including that Personal Data shall be processed lawfully, fairly and in a transparent manner that ensures the security of the Personal Data. It is initially proposed that only the Lab research team will have access to data collected as part of the evaluation. However, it is expected that in order to review the fidelity of the implementation of GenPMTO, and to coach and certify practitioners, ISII will require access to data from the sessions in which the practitioners deliver the programme to caregivers. Given that most ISII training staff are based in the USA, Barnardo's have established a Data Sharing Agreement with ISII, with the Lab overseeing this process.

For the duration of the evaluation, the Lab will be the data controller who also processes data. This means that the Lab is responsible for deciding the purpose and legal basis for processing data. The legal basis is "legitimate interest". Article 6(1)(f) of UK GDPR states that "*processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child*."

The Lab has determined there is a genuine purpose to process this data. This data will inform the necessary evidence around what works to improve caregivers' skills and strategies, which in turn, improves positive behaviours, relationships and life outcomes of young people, particularly those at risk of, or who have engaged in, violent behaviours. Data processing is necessary to complete a robust evaluation. The Lab does not consider that collecting and gathering data for this trial will interfere with individuals' interests, rights or freedoms. The data subjects will include; at-risk youth, caregivers of at-risk youth, the developer team at ISII and the delivery team at Barnardo's.

During this trial, data will be stored on secure, password-protected and encrypted network drives (hosted by BIT). Access to the data will be restricted to the relevant members of the project team involved in this evaluation.

All data shared with BIT will be processed in line with its data protection policy. A summary of this policy can be found in Annex C. In the analysis, BIT will promote data quality and security through the following measures.

- All variables will be clearly named, coded and labelled before analysis
- Checks on the data received will be carried out for valid values, range, and consistency against already held data
- Any modifications to datasets will be recorded in the analysis code, which will be well-annotated
- Original raw datasets will never be amended
- Access to the project data will be restricted to project personnel
- All data stored by BIT will be backed up

In case a Personal Data Breach occurs despite the mitigations in place, project team staff will deal with the security incident without undue delays. All Personal Data Breaches (or suspected Personal Data Breaches) will be reported to BIT's Data Protection Officer as soon as a project team member becomes aware of one (including if this is outside of office hours) by contacting the Data Protection Officer directly and by completing a Data Incident Notification Form. Staff will not attempt to investigate a Personal Data Breach themselves but will take steps to contain the Personal Data Breach as quickly as possible. Such steps might be taken prior to reporting the incident to the Data Protection Officer where this is reasonable and necessary to protect Data Subjects and mitigate the potential impact of the Personal Data Breach.

Data management

All quantitative and qualitative data will be stored in a secure Google Folder where access is restricted to only researchers conducting the analysis. Data will be deleted upon completion of the project 6 months after the conclusion of the project.

Quantitative data

Survey data

We will use SmartSurvey to collect the survey data. SmartSurvey produces a spreadsheet where one row is a survey response. This will be used to code the survey outcomes using the methods outlined in the outcome measures table.

Surveys will ask participants to record their name. This enables us to link survey responses with demographic data and other outcome measures. Once survey responses have been linked, participants' names will be removed.

Programme administrative data

Barnardo's is responsible for providing us with the programme administrative data. All data shared with the lab by Barnardo's will be received via a secure transfer link (Virtru or Quatrix).

Programme administrative data includes the *referral data* and the *programme delivery data* (e.g. attendance sheets, fidelity checklists). *Referral data* will be collected via an online form (located on FormAssembly) completed by borough staff. Barnardo's will download the data in a spreadsheet and share the relevant data with the lab. *Programme delivery data* will be collected via Barnardo's , and shared with the lab.

Qualitative data

Interview transcripts

Interview recordings will be uploaded to McGowan for transcription. All interview recordings will be transcribed verbatim to ensure accuracy and facilitate subsequent analysis. Transcripts will be anonymised by assigning unique identifiers to each participant, replacing their names or any identifying information. Transcripts, observation notes, and any additional relevant documents will be securely stored in a password-protected file area. Access to the data will be restricted to only project team members involved in the analysis. Recordings will be deleted upon completion of the project.

9. Racial diversity and inclusion

The Lab is committed to conducting research in which equality, diversity and inclusion principles are firmly embedded across all stages of evaluation, from the design, recruitment, data collection, and analysis.

Groups included in the programme and evaluation
The Lab will work with Barnardo's to monitor for inequalities within the referral and recruitment processes to ensure that no demographic group is unduly excluded from access to the study and the GenPMTO programme. At all stages, we will monitor whether certain demographic groups are under- or over-represented in referrals to the programme by referral agencies. This may occur due to unconscious bias within referral agencies, and/or because the study or GenPMTO programme is viewed by referral agencies as unsuitable for families with certain demographic characteristics. In either case, the Lab and Barnardo's would seek to investigate this further in consultation with referral agencies.

Similarly, at all stages the Lab will work with Barnardo's to monitor whether the rate of acceptance to the trial (i.e. families accepting the offer to participate in the RCT) varies across certain demographic groups. If this is the case, the Lab and Barnardo's will seek to investigate why this is the case, and whether the programme content and/or delivery needs to be adapted to ensure equality of acceptability and access.

Inclusivity during recruitment and programme delivery
The Lab will work with Barnardo's as programme delivery partner to ensure that inclusive practices are central to the *recruitment process* and that participant wellbeing is promoted by:

1. Being considerate of the sensitivity of the topic area during recruitment
2. Providing caregivers with welcoming information documentation, which provides all necessary information about data security, anonymity and the reasons for undertaking research, in plain English
3. Offering a flexible and varied range of times for introductory (and other) sessions, and the option of attending sessions remotely via video-link, allowing different groups and individuals the opportunity to participate.

Inclusivity during data collection
The collection of data directly from caregivers will occur via caregiver surveys and caregiver interviews. To ensure that the principle of inclusivity is adhered to during this process, the Lab will work with Barnardo's to:

1. Use inclusive and accessible language in all survey and interview questions and guidance;
2. Ensure that sufficient numbers of either Lab and/or Barnardo's staff will be present at in-person sessions where caregiver surveys are being administered, to provide assistance or instruction as required;
3. Strive for equality of access by enabling online (remote) participation in caregiver surveys and interviews. Access issues could include a lack of time during the day to attend sessions, or distance from an in-person session.
4. Training for researchers: Prior to conducting interviews, researchers will complete the NSPCC's Introduction to safeguarding and child protection training.

Wellbeing and safety during surveys and interviews

The Lab is conscious that families who engage in the evaluation could be vulnerable to negative and stressful impacts of the research process. The Lab will work to ensure the wellbeing and psychological safety of individuals during data collection by:

1. **Designing interview questions to minimise harm and maximise comfort:** The Lab will maximise wellbeing and minimise harm during surveys and interviews by (i) structuring questions to build in complexity and difficulty to increase comfort as rapport to develop, (ii) depersonalising questions to elicit comfort and stronger answers (e.g. instead of 'what do you hate about X', ask 'If you had a magic wand, what 3 things would you change about X?"), (iii) being aware of tension discomfort or distress during the interview, repeating that the interview can be stopped may help participants and repeatedly ask if they want to continue, (iv) ensuring that researchers are aware of places to signpost participants and offer this information, and (v) auditing the questions for their sensitivity within the context before the interview.
2. **Allowing the participants to choose their environment for participating**: Where possible the Lab will allow the interviewees to make decisions about the survey and interview setting(s) - at their home, a public place or over the phone, enabled by the online conference format.
3. **Reminding participants of anonymity and data security**: The Lab will seek to minimise anxiety for caregivers by reminding them that the process is fully anonymous and that all identifiable information will be removed from the transcripts and report.

10. Risks

We have identified the following risks, focusing specifically on the evaluation:

Risk	Impact	Likelihood	Mitigation
Low participation in research activities by caregivers (including differential attrition and higher drop-out from the control group)	High	Medium	• Providing incentives to the control and treatment groups (£20 per completed survey); • Ensuring that evaluation activities are designed to be low-impact in terms of burden and time (including offering flexibility in terms of how surveys are completed); • Where needed, utilising the relationships that Barnardo's and referring practitioners have built with participants to facilitate access and cooperation. • We will inform participants with sufficient notice about planned survey activities. In particular we will attempt to maintain contact throughout the trial period particularly with control group participants, to remind them of their participation in the study and upcoming opportunities to receive their vouchers. • We will monitor rates of attrition, and whether there are differences in rates of attrition based on parent/caregiver ethnicity, gender, age, and

			in terms of baseline outcomes.
Unrepresentative business-as-usual offer	High	Medium	• Given variation in the level of need of families referred into the project, and variation in how, why, and by whom they've been referred, we would expect the business-as-usual for these families to naturally involve a mix of services, and for the business-as-usual offer not to be dominated by high-intensity programmes similar to GenPMTO. • We have emphasised in all communications with Boroughs that we do not want control groups to receive an artificially restricted or artificially boosted offer in the context of this trial, and that control group participants should receive services that the Borough would have provided to these caregivers in the absence of the GenPMTO project. • We will monitor what services the control group receive.
Non-compliance by staff with random assignment	High	Low	• Our delivery team and Boroughs understand and are on board with the requirements of running a randomised controlled trial. • We will monitor compliance through attendance lists to identify any accidental crossover from control group to treatment group.
Spillover between treatment and control participants	Medium	Low	• The likelihood of spillover is deemed to be low given that the programme is covering a large geographical area with over 200,000 residents. • Moreover, the programme is intensive and involves a variety of roleplay exercises and activities to support parents in developing and implementing new parenting strategies. We believe that any effects are likely to emerge as a result of participation in the full programme, and don't expect that sharing information from the programme would provide the same benefits. • Finally, if spillover does occur, we believe this would lead to an improvement in outcomes for the control group, which would likely cause an underestimation of the effect size of the intervention, which is relatively tolerable as a risk (when compared to an overestimation).
There is turnover at the Lab, such that we do not have sufficient capacity to collect and analyse data, and manage the overall project	Medium	Low	• Should core members of the evaluation team be taken ill, or depart, EYV Lab is able to quickly draw on the resource pool from BIT, which consists of over 150 staff members. • Key decisions and files are stored in secure folders on Google drive. These will help team members to aid with knowledge transfer and

			onboarding of new staff, should other core EYV Lab staff depart. ● Thorough trial documentation (including this protocol) will allow for the trial to be evaluated and the results analysed, even if key staff are redeployed or depart.

11. Timeline

Activity	Cohort	Target dates	Description
Recruitment	All cohorts.	November ‘24 to June ‘25	Rolling recruitment over a 7 month period.
Baseline data collection and randomisation	1	December ‘24	-
	2	January ‘25	
	3	February ‘25	
	4	March ‘25	
	5	April ‘25	
	6	May ‘25	
	7	June ‘25	
GenPMTO delivery	1	December ‘24 to March ‘25	-
	2	January ‘25 to April ‘25	
	3	February ‘25 to May ‘25	
	4	March ‘25 to June ‘25	
	5	April ‘25 to July ‘25	
	6	May ‘25 to August ‘25	
	7	June ‘25 to	

		September ‘25	
Post-test data collection	1	March ‘25	-
	2	April ‘25	
	3	May ‘25	
	4	June ‘25	
	5	July ‘25	
	6	August ‘25	
	7	September ‘25	
Analysis	-	March ‘25 to October ‘25	Ongoing data management and analysis throughout study as cohorts begin to end.
Reporting and progression decision to stage 4 efficacy	-	October ‘25	-
6-month follow-up data collection	1	September ‘25	-
	2	October ‘25	
	3	November ‘25	
	4	December ‘25	
	5	January ‘26	
	6	February ‘26	
	7	March ‘26	

NB: We will submit data to YEF’s archive at the point of study completion.

Annexes

Annex A: Summary of GenPMTO programme using the TIDieR framework

Name: Provide a name or phrase that describes the intervention.	GenPMTO
Why: Describe any rationale, theory, or goal of the elements essential to the intervention.	• Positive parenting practices promote positive child/youth outcomes, and coercive parenting practices disrupt them. As children become adolescents, peers also become mediators of youth outcomes. • The core positive parenting practices are: skill encouragement, limit setting, monitoring/supervision, family problem solving, and positive involvement. • The programme aims to teach parents effective parenting strategies, increase effective parenting, and reduce deviant peer association, which then mediate programme effects on positive child/youth outcomes. • In the short term, the programme aims to reduce children and young people's internalising and externalising behaviour problems. • In the longer term, the programme aims to reduce police arrests, increase school functioning, improve social relationships, and reduce substance use. • Parents show improved marital relationships, a rise out of poverty, and increased socio-economic status.
What - Materials: Describe any physical or informational materials used in the intervention, including those provided to participants or used in intervention delivery or in training of intervention providers. Provide information on where the materials can be accessed.	• Training materials and implementation guides for practitioners. • Materials for parents/caregivers. • Materials are purchasable from ISII.
What - Procedures: Describe each of the procedures, activities, and/or processes used in the intervention, including any enabling or support activities.	• **Initial training and technical assistance -** (ISII) provides extensive training and coaching to local practitioners, some of whom are eventually certified as leaders, mentors, trainers, coaches, or fidelity raters for the following generations of PMTO clinicians. • **Ongoing fidelity monitoring -** Practitioners videotape programme sessions for review and coaching by ISII.

	Practitioners upload written materials and video recordings to ISII's online platform, and receive feedback. ● **10- to 14-session parenting programme -** Practitioners use active teaching approaches (such as group problem-solving, role-play, and video modelling) to support caregivers in using positive parenting strategies at home.
Who: For each category of intervention provider (such as psychologist, nursing assistant), describe their expertise, background, and any specific training given.	Qualifications required for practitioners depend on the agencies that employ them. Practitioners may have Bachelor's, Master's, or Doctorate level degrees. Practitioners serve in a wide variety of delivery systems including child welfare, youth justice, and child mental health.
How: Describe the modes of delivery (such as face to face or by some other mechanism such as internet or telephone) of the intervention and whether it was provided individually or in a group.	The intervention will be provided in group-format, either face-to-face or online. More broadly, the intervention can be delivered on an individual basis, but this mode of delivery will not be used as part of this project.
Where: Describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features.	The programme can be delivered in out-patient health settings, the home, and community centres/settings. In this project, the programme will be delivered in community centres and settings.
When and how much: Describe the number of times the intervention was delivered and over what period of time including the number of sessions, their schedule, and their duration, intensity, or dose.	The programme can be delivered over different modes (i.e. in-person or online). Sessions are between 90 and 120 minutes each.
Tailoring: If the intervention was planned to be personalised, titrated or adapted, then describe what, why, when and how.	Programme adaptations made as part of the Stage 1 study. These will be reported on in future publications.
Modification: If the intervention was modified during the course of the study, describe the changes (what, why, when, and how).	To be assessed on an ongoing basis as part of the study.

How well (planned): If adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them.	Programme developer experts assess practitioner sessions using the FIMP (Fidelity of Implementation Rating System) measure, developed by ISII.
How well (actual): If actual adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned.	To be assessed on an ongoing basis as part of the study.

Annex B: Power calculations

Power calculations using a within-group design can be found below.

Sample sizes required for MDES on young people's externalising behaviour/conduct problems, accounting for inclusion of pre-test covariate[29]

Between-group trial		
n	***n* (per arm)**	**MDES (Cohen's d)**
100	50	**0.48**
150	75	**0.43**
200	100	**0.34**
240	120	**0.31**
250	125	**0.30**
300	150	**0.27**
400	200	**0.24**
500	250	**0.21**
600	300	**0.19**

Assumptions

Power calculation was determined using the following assumptions:

- **Assumed power:** 0.8
- **Significant level:** 0.05
- **Level of randomisation:** individual-level (i.e. individual families will be randomised)
- **Number of trial arms:** 2 (i.e. intervention and control)
- **Outcome measure:** child externalising behaviour/conduct, as measured by the Eyberg Child Behavior Inventory (ECBI)
- **Estimated standard deviation**: 9.42
- **Estimated mean:** 52.02
- **Estimated pre-post correlation:** 0.53

Sample sizes required for MDES on parenting outcomes, without the inclusion of pre-test covariate

Between-group trial		
n	***n* (per arm)**	**MDES (Cohen's D)**
100	50	**0.56**

[29] Using pre-post correlations reported in Bjørknes & Manger 2012.

150	75	**0.46**
200	100	**0.39**
240	120	**0.36**
250	125	**0.35**
300	150	**0.33**
400	200	**0.28**
500	250	**0.25**

Assumptions

Power calculation was determined using the following assumptions:

- **Assumed power:** 0.8
- **Significant level:** 0.05
- **Level of randomisation:** individual-level (i.e. individual families will be randomised)
- **Number of trial arms:** 2 (i.e. intervention and control)
- **Outcome measure:** Parenting practices measured with the Parenting Practices Interview (PPI)
- **Estimated standard deviation**: 1.41
- **Estimated mean:** 67.51

Annex C: BIT data protection policy summary

The General Data Protection Regulation (GDPR) imposes certain obligations upon Behavioural Insights Limited (BIT), and other companies within the group, as Controllers and / or Processors in relation to processing Personal Data.

BIT takes these obligations seriously. BIT is committed to respecting the rights of all individuals whose personal data it processes:

1. **In relation to data security,** BIT has implemented appropriate measures to ensure the secure storage and handling of Personal Data, including obtaining a Cyber Essentials Plus certification and developing a comprehensive Data Handling Protocol.
2. **In relation to data protection and privacy rights,** our data processing activities are conducted according to the principles relating to the processing of Personal Data set out in the GDPR, including that Personal Data shall be processed lawfully, fairly and in a transparent manner, and in a manner that ensures the security of the Personal Data. BIT has policies and procedures in place to ensure compliance with these principles.

More information on how we handle Personal Data in relation to projects we are working on is detailed below.

BIT is registered with the UK ICO under the terms of the Data Protection Act 2018. BIT's registration number is ZA038649.

Privacy by design

BIT conducts all trials and research projects with a privacy by design approach to protect and maintain the privacy and security of research participants' and research subjects' data. We work closely with clients, government departments and research partners when designing interventions to ensure that a privacy by design approach is implemented and respected.

Our data protection and data security policies and procedures reflect necessary legislative requirements and set out the standard to which BIT staff should work when dealing with Personal Data, including:

- Attendance at mandatory data protection training for all employees;
- Identifying data requirements from the outset of each project;
- Minimising use of Personal Data where possible and ensuring we have the right to handle any Personal Data where successful project delivery is reliant on using it;
- Putting in place data processing agreements with all clients and suppliers to clarify data handling arrangements ahead of any data being transferred;
- Complying with all relevant data residency requirements and implementing appropriate technical and organisational measures, to protect data and avoid unauthorised access, internally and externally;
- A clear internal reporting process in the event of a data breach, to consider the nature of the breach and identify any necessary action, including whether the breach should be reported to the relevant authorities, i.e. the Information Commissioner's Office in the UK or the Office of the Australian Information Commissioner;
- Clear procedures on retention and destruction of Personal Data to avoid keeping hold of Personal Data longer than necessary for the purposes of each project; and
- Implementing robust investigation and reporting procedures in relation to any data breach or security issues that arise both within our own systems and those of our clients, partners and suppliers.

Data Protection Officer

The BIT group of companies has appointed a Data Protection Officer (DPO) who is the first point of contact for any issue regarding data protection and data security. The DPO can be contacted via email at dpo@bi.team or by writing to us at: Data Protection Officer, Behavioural Insights Limited, 58 Victoria Embankment, London, EC4Y 0DS, United Kingdom.

Annex D: Evaluation team experience

- **Tom McBride** is the Director of the Ending Youth Violence Lab and has over 15 years of experience in research and evaluation roles. He is the former Director of Evidence at the Early Intervention Foundation and Head of Strategic Analysis at the Department for Education. Tom will have overall responsibility for the delivery and quality of this work

- **Jack Martin** is an Assistant Director within the Ending Youth Violence Lab and has over 8 years of experience working at the Early Intervention Foundation and sits on the Government's Trials Advice Panel. Jack will oversee the delivery of the work and support, supervise and quality assure the work of the project team.

- **Patrick Taylor** is a Principal Research Advisor and leads BIT's education and youth evaluation work, supporting the design, improvement and evaluation of complex interventions in these fields. Patrick will provide support and quality assurance for the pilot evaluation.

- **Lilli Wagstaff** is a quantitative research advisor in the Home Affairs and Security team at BIT and leads the evaluation and day-to-day delivery of a number of projects focusing on policy areas including reducing violence and recidivism. Lilli will lead the quantitative evaluation.

- **Niall Daly** is a Research Advisor in the Health and Wellbeing team at BIT, specialising in trial design, implementation, and quantitative data analysis across a range of projects within the health space. He will support the quantitative evaluation and broader project management of the work.

- **Emma Forsyth** is an experienced mixed methods social researcher with a demonstrated history of working in applied policy research. Emma will support the qualitative research.

- **Ivana La Valle** is a research consultant with extensive experience of carrying out research to inform and evaluate children's policy and practice. Ivana hasI led a number of large scale national studies that have played a key role in shaping children and families policy in the past 15 years. Ivana will lead the qualitative research.

- **Dr Sajid Humayun** is a senior lecturer in psychology at the University of Greenwich. Sajid is an expert in youth justice and in evaluating interventions for youth crime. Sajid worked on the first RCT for a County Lines intervention and ran the first British evaluation of Functional Family Therapy. Sajid will be providing expert advice and challenge on the design and delivery of the evaluation on a consultancy basis.