EVALUATION REPORT

Tavistock Relationships MBT-PP

Pilot trial report

Abigail Millings, Elaine Clarke, Sean Demack, John Reidy, Madelynne Arden, Charlotte Coleman, Laurynas Rutkauskas, Piyali Misquitta, & Anna Stevens

January 2025

Sheffield Hallam University | Centre for Behavioural Science and Applied Psychology

YOUTH ENDOWMENT FUND

About the Youth Endowment Fund

The Youth Endowment Fund (YEF) is a charity with a mission that matters. We exist to prevent children and young people from becoming involved in violence. We do this by finding out what works and building a movement to put this knowledge into practice.

Children and young people at risk of becoming involved in violence deserve services that give them the best chance of a positive future. To make sure that happens, we'll fund promising projects and then use the very best evaluation to find out what works. Just as we benefit from robust trials in medicine, young people deserve support grounded in the evidence. We'll build that knowledge through our various grant rounds and funding activities.

And just as important is understanding children and young people's lives. Through our Youth Advisory Board and national network of peer researchers, we'll ensure young people influence our work and that we understand and are addressing their needs. But none of this will make a difference if all we do is produce reports that stay on a shelf.

Together, we need to look at the evidence and agree on what works, then build a movement to make sure that young people get the very best support possible. Our strategy sets out how we'll do it. At its heart, it says that we will fund good work, find what works and work for change. You can read it here.

For more information about the YEF or this report, please contact:

Youth Endowment Fund
C/O Impetus
10 Queen Street Place
London
EC4R 1AG

www.youthendowmentfund.org.uk

hello@youthendowmentfund.org.uk

Registered Charity Number: 1185413

Contents

About the evaluator

This evaluation is being conducted by the Centre for Behavioural Science and Applied Psychology (CeBSAP) at Sheffield Hallam University. The evaluation team comprises:

Prof. Abigail Millings, Professor of Applied Social Psychology and Project Lead

Dr Elaine Clarke, Research Fellow and Trial Manager

Sean Demack, Deputy Head of Sheffield Institute of Education Research & Knowledge Exchange and Statistician

Prof. John Reidy, Emeritus Professor and Statistician

Prof. Madelynne Arden, Professor of Health Psychology, Director of CeBSAP and Treatment Fidelity Lead

Dr Charlotte Coleman, Deputy Head of Psychology, Sociology and Politics and Implementation and Process Evaluation Advisor

Dr Laurynas Rutkauskas, Researcher

Piyali Misquitta, PhD Candidate

Anna Stevens, Research Fellow and Data Management Advisor

Contact email: a.millings@shu.ac.uk / cebsap@shu.ac.uk

Executive summary

The project

Mentalization Based Therapy for Parenting Under Pressure (MBT-PP) is a brief, manualised course of psychotherapy which aims to reduce interparental conflict (IPC), whether parents are in a relationship or separated. Although only parents participate in the therapy, reduced parental conflict is hypothesised to reduce the social and emotional difficulties their children may experience. Developed and delivered by Tavistock Relationships, MBT-PP is delivered for up to 16 weeks to parents of children aged 8–14 who are experiencing intense and ongoing IPC. The programme consists of ten therapy sessions delivered via online video-call by MBT-PP practitioners who are qualified therapists. After individual assessment sessions with each parent, both parents take part in eight joint sessions with the therapist. If only one parent is willing to participate, the therapist works with that parent on their own while seeking to engage with the second parent for future sessions. The joint sessions are dedicated to improving parents' capacity to 'mentalise' or make sense of their own and others' thoughts, beliefs and emotions in real time, with the goal of equipping them to consider their partner's and child's feelings during emotionally intense moments.

The Youth Endowment Fund (YEF) funded an internal pilot study as the first phase of an efficacy trial of MBT-PP. The pilot phase sought to test out the research processes involved in conducting a randomised controlled trial of the programme. This included examining how well the referral pathways, data collection and data linking processes worked and whether the processes required for a trial were acceptable to families. It also aimed to assess whether the drop-out rate was too high, estimate how many families would need to be recruited for a full efficacy trial and explore whether there were very early indications to support the theory that MBT-PP reduces parents' conflict and improves children's social and emotional difficulties. To be eligible for the study, parents either had to have been referred by local authority staff or, if they self-referred, had to be classified as a distressed couple by one or both of their scores on the Dyadic Adjustment Scale (a measure of relationship quality). Families were then randomised to receive either MBT-PP or treatment as usual (TAU). Interviews were conducted with the three local authority referral leads, who were responsible for recruiting families and coordinating the project. Parents and children completed several quantitative surveys before the therapy started and after it was complete, including the parental-report version of the Strengths and Difficulties Questionnaire (SDQ), a measure of children's social and emotional difficulties. During the pilot phase, which ran from June 2023 to April 2024, 149 parents of 132 children in three local authorities (Bristol; Dorset; and Bournemouth, Christchurch and Poole) were recruited and randomised. The majority of children (94%) were identified as White, 5% as Mixed, 1% as Asian, and 1% as Other. At the end of the pilot, post-intervention SDQ data were available for 45 children, with most still undergoing treatment (MBT vs TAU).

Key conclusions
Setting up the processes required to recruit families into the study was time-consuming. Local authority referral leads reported that referral pathways took a long time to create and embed. Although these pathways are now reported to be working, recruitment remains a time-intensive task.
The research processes (including consent, eligibility screening and randomisation) were broadly acceptable to families. Most families that had been referred to the programme agreed to be screened for eligibility, although 29% declined. None of the families that had been deemed eligible dropped out before they were randomised, indicating that randomisation is acceptable.
MBT-PP was well attended, with 89% of families that started the programme attending six or more sessions. However, 22% of families dropped out before completing the post-intervention surveys. As these surveys include the primary outcome measure (the mother's SDQ score), this may pose a problem for the project moving forward.
When parents did fill out surveys, the quality of the data was high. The research team was able to effectively link individuals' responses at baseline and follow-up using unique family and participant ID numbers.
The evaluator judges that a larger, randomised controlled trial is feasible and estimates that 250 families (including 350 children) would be required for a robust impact trial.

Interpretation

Generating the local knowledge required to produce a steady stream of referrals into the project was time-consuming. Local authority referral leads reported that referral pathways took a long time to create and embed. Promotional materials for the study were not co-designed, and local authority referral leads reported that their wording may have made it difficult to recruit parents and establish relationships with organisations that could provide referrals. Nevertheless, recruitment numbers increased steadily over time, and the proportion of eligible referrals also improved and stabilised, with roughly 50% of referred families being eligible to participate in the study, 20% being ineligible and 30% disengaging. However, only 90 families from the target of 140 were recruited for the pilot. Although these referral pathways are now reported to be working, recruitment remains a challenging and laborious process. Local authority referral leads explained that onboarding families takes longer than expected due to the volume of administration required.

The research processes required for a trial (including consent, eligibility screening and randomisation) were broadly acceptable to families. Most families that had been referred to the programme agreed to be screened for eligibility, although 29% declined. Local authority referral leads consistently reported that making prompt contact with referred parents was key to recruitment. They also reflected that the reference to parental conflict in the programme branding was a barrier for many parents. A rebranding is underway, which puts the focus on improving communication between parents, to improve parents' willingness to proceed with the programme. None of the families that had been deemed eligible dropped out before they were randomised, indicating that randomisation is acceptable.

MBT-PP was well attended, with 89% of families that started the programme attending six or more sessions. However, 22% of families dropped out before completing the post-intervention surveys (eight families in total, four from each arm of the study). This poses a potential problem, as the primary outcome measure, the mother's SDQ score, is collected at this point. The evaluators plan to improve post-intervention survey completion by restructuring the voucher incentives and preparing survey invitation emails informed by behavioural science.

When parents did fill in surveys, the quality of the data provided was high. Only 0.04% of questions on the baseline SDQ survey went unanswered, and none were missed at post-intervention. Participants' answers to similar questions on the same survey were generally consistent, suggesting that the surveys were reliable. The correlation between co-parents' judgements of their children's difficulties on the SDQ was moderate at baseline and strong post-intervention, possibly reflecting greater agreement after receiving MBT-PP. However, parents' alignment on the emotional symptoms subscale decreased post-intervention. Given that high levels of parental conflict could influence how aligned parents' views are, the evaluators decided not to substitute the father's SDQ score when the mother's score is missing.

The research team was able to effectively link individuals' responses at baseline and post-intervention using unique family and participant ID numbers. As an initial exploration of the potential impact of MBT-PP, the evaluators examined changes in child and parent outcomes post-intervention. These were mixed and unclear: several improved, others declined, and patterns of change varied between mothers and fathers. However, the sample size for the pilot was small, meaning that no conclusions can be drawn from this.

The data from the pilot phase led to adjustments in the estimate of the number of families that would be needed for the full efficacy trial. The initial power calculation indicated that a sample of 350 families (including 700 children) would be required. The updated calculation reduced the recommended sample to 250 families, including 350 children. This sample size would allow for 20% of families to drop out of the evaluation. The estimated rate of 17 referrals per month means that to meet this sample size, recruitment for the full trial would need to be extended by an additional three to six months.

The YEF is proceeding with the efficacy trial of MBT-PP, to which this pilot contributes. This is due to report in 2026.

Introduction

Background

Interparental conflict (IPC), whether in separated or intact families, is consistently related to poorer child adjustment (van Eldik et al., 2020), and this link is often mediated by parenting style (van Dijk et al., 2020). Frequent, intense and unresolved conflicts between parents significantly impact children's mental health across various domains. These include sleep disturbances; externalising problems, such as aggression and antisocial behaviour; internalising issues, such as anxiety and depression; academic difficulties; social and interpersonal relationship challenges; and physical health problems (Harold and Sellers, 2018).

Minority ethnic families are disproportionately affected by conditions such as poverty, which increases their vulnerability to stressors. Research shows that ethnic minorities experience higher levels of economic hardship and stress, which in turn magnifies the risk of IPC and disrupted parenting practices (Masarik and Conger, 2017). A report from the Department for Work and Pensions (DWP, 2017) highlighted that among families experiencing poverty, around one-third of all children in workless couple-parent households lived with parents in distressed relationships.

Research suggests that children fare better following parental separation or divorce when their parents engage in supportive and cooperative co-parenting and that the absence of this can be a risk factor for poor child outcomes, such as emotional and behavioural problems and poor academic outcomes (Adamsons and Pasley, 2006). IPC in both intact and separated families is problematic for children, with hostility being related to externalising behaviours and emotional responses (van Eldik et al., 2020). However, it is important to note that low relationship quality, conflict frequency and, specifically, child-related conflict in parental relationships are also damaging (van Eldik et al., 2020).

There are multiple models and theories that focus on the mechanisms by which IPC produces negative impacts on children and young people. The interconnected nature of family processes means that children's well-being is affected by the interactions between IPC, parent–child relationships and, crucially, children's interpretations of these conflicts. In an extension of the Family Stress Model (Conger et al., 1994), Acquah et al. (2017) highlighted as central the additional strain caused by economic pressure. Acquah et al. (2017) propose that economic pressure increases each parent's distress, which in turn precipitates or increases conflict. IPC then has both a direct effect on child outcomes and an indirect effect on child outcomes via the parent–child relationship.

Harold et al. (2016) identified three primary theoretical models, the Cognitive-Contextual Framework (Grych and Fincham, 1990), the Emotional Security Hypothesis (Davies and Cummings, 1994) and the Family-Wide Model (Harold and Conger, 1997), that converge on the idea that children's subjective understanding of IPC is crucial in explaining variations in their psychological responses. Acquah et al. (2017), therefore, also include child appraisals of IPC as a factor that can moderate the effects of IPC on parent–child problems and child outcomes. Aquah et al.'s (2017) model explains that the way in which children perceive and contextualise IPC determines the extent to which such conflict impacts their psychological development. For example, children who can attribute the conflict to external factors such as economic stress, rather than to themselves, may be more resilient and better able to maintain quality relationships with their parents.

These models, which are integrated within the Process Model of Family Stress on Children's Mental Health Problems, Including Child Appraisals of Inter-Parental Conflict (Acquah et al., 2017), highlight that children's evaluations of conflict influence their emotional security, their expectations in family relationships and their long-term psychological health.

While the short-, medium- and long-term damaging effects of IPC on children are well established, there is a paucity of evidence-based interventions to reduce IPC. The Process Model of Family Stress on Children's Mental Health Problems, Including Child Appraisals of Inter-Parental Conflict (Acquah et al., 2017) indicates that such interventions should focus on enhancing children's understanding of the conflict, coaching parents on addressing their children's concerns and fostering supportive parenting strategies. By doing so, we can better support children in coping with familial stressors and reduce the risk of long-term emotional and behavioural issues. Mentalization Based Therapy for Parenting Under Pressure (MBT-PP) has the potential to achieve these goals, but it lacks the evidence base required for a wide-scale roll-out.

An additional barrier is the lack of an established referral pathway for families experiencing IPC to access support. Because IPC is a problem that crosses multiple domains, such as health, education, social care, family law and economic security, there is no obvious single referral point, and a high degree of coordination between sectors would be required to create one. That said, significant inroads have been made in recent years to tackle the issue. Two notable examples come from the DWP and the Ministry of Justice (MoJ). The DWP's Reducing Parental Conflict (RPC) programme funded the delivery of IPC interventions in 31 upper-tier local authorities (LAs) from 2018 to 2022. The RPC evaluation illustrated the scale of the challenge of creating referral pathways and highlighted that 'prior to being approached by the RPC programme, it was common for local authorities not to have thought about tackling parental conflict below levels amounting to domestic abuse. In many areas, parental conflict had not historically been seen as a policy area or priority' (DWP, 2020). Separately, in early 2021, the MoJ announced two pathfinder pilot court sites that have the authority to pilot new ways of working with family separation to combat both the family court demand crisis and the propensity for many troubled families to seek legal solutions to problems that would benefit more from therapeutic support (likely because no such therapeutic offering exists). Heralded a success, the pathfinder pilot project has been expanded to include more sites ahead of a national rollout. This policy landscape illustrates both the need and political will to create better access to appropriate support for IPC and highlights the urgency of establishing an evidence base for IPC interventions.

Mentalization Based Therapy (MBT, Bateman and Fonagy, 2010) is a psychotherapeutic approach that involves working on the ability to 'mentalise', or think about thinking processes in real time. It helps people to make sense of their thoughts and feeling states and to also be able to apply these skills to thinking about the reasons for other people's behaviours. Research suggests that MBT may be effective for borderline personality disorder (BPD), self-harm in adolescents and mothers involved in substance abuse treatment (Malda-Castillo, Browne and Perez-Algorta, 2018). Reviews are cautiously supportive, with the caveat that the evidence base is limited and requires more rigorous studies (Malda-Castillo, Browne and Perez-Algorta, 2018; Vogt and Norman, 2018). A meta-analysis found that while MBT was promising for reducing BPD, depression and self-harm symptoms, it was no more efficacious than control conditions (Hajek Gross et al., 2024), yet systematic reviews point to the notion that the ability to mentalise is a key factor in psychotherapeutic change (Lüdemann, Rabung and Andreas, 2021; Luyten et al., 2024).

Due to its focus on understanding one's own and others' thoughts and feelings in difficult interactions, MBT is particularly suitable to the context of the impact of IPC on children and young people. MBT-PP is a brief, manualised therapy programme designed to reduce IPC by supporting the parents to mentalise about their

child's experience and each other's motivations and experiences. It was first piloted in a small-scale study (Hertzmann et al., 2016, described below). The MBT-PP intervention was subsequently refined by the developer for the DWP's RPC programme, and it was rolled out as part of that programme in 2022, utilising online delivery (face-to-face therapy delivered via video call). By developing the capacity to mentalise about the other people involved in the conflict, parents learn to understand their child's perspective and the impact the conflict has on them. Parents also learn to understand each other's perspectives, such that actions which might previously have been attributed to hostile motivations are viewed in a more balanced way. The insights that parents glean into each other's and their child's experiences, particularly around conflictual interactions, lead to a reduction in conflict and more adaptive ways to manage disagreements. A small-scale random allocation feasibility study compared MBT-PP to treatment as usual (TAU; a parents' group) and found encouraging support for this (Hertzmann et al., 2016; 2017). Thirty parents (15 pairs of separated co-parents who were entrenched in chronic and intense conflict over their children) completed quantitative measures and qualitative interviews (pre- and post-intervention) to explore several outcome variables. Both intervention groups (i.e. the MBT-PP group and the TAU group) showed statistically significant improvements in reported 1) expressions of anger towards the ex-partner, 2) levels of stress and depression and 3) behavioural and emotional difficulties experienced by their children. Furthermore, in both groups, attitudes towards the ex-partner improved. Following MBT-PP in particular, parents' descriptions of their ex-partners became less polarised, and they were more able to accept that their co-parent was likely experiencing similar feelings and motivations to those they were experiencing themselves (Hertzmann et al., 2016; 2017). Importantly, while the study failed to detect a significant difference in the parents' abilities to mentalise according to the quantitative measures, the qualitative findings suggested that nuanced shifts in this had occurred. The authors posited that the study may only have detected the first part of a process of change, with a larger, longer-term study better able to establish what was going on (Hertzmann et al., 2016). Overall, the study showed promise and suggested that a full-scale randomised controlled trial (RCT) was warranted. MBT-PP was provided to over 1,000 parents in 2022 under the DWP's RPC programme (DWP, 2021; 2023), further establishing the feasibility of the intervention.

Internal pilots enable data from the pilot and efficacy stages to be combined in order to increase sample size (and hence statistical sensitivity). Internal pilots are best thought of as smaller-scale efficacy trials which might be undertaken to gain some evidence of promise before committing funding to larger-scale efficacy trials. All types of pilot could provide evidence of promise and are also useful for obtaining empirical estimates to help improve the precision or accuracy of power analyses for an efficacy trial. What makes 'internal pilots' distinct is the pre-specified plan to combine data from the pilot and efficacy stages. This results in greater restrictions on adaptations between these stages, most strongly around the intervention (e.g. how it is implemented and theorised) but also around evaluation design (primary outcome, trial design). Programmes that are well developed and have some exposure to type-3 evaluation methodologies (quasi-experimental design or RCT) are considered suitable for internal pilots. MBT-PP meets the criterion of being well established, and evidence of promise was found in a small-scale, underpowered RCT (e.g. Hertzmann et al., 2016; 2017) and in the RPC programme (DWP, 2023). The current design, therefore, assumes the suitability of an efficacy study with an internal pilot. This report presents the findings from the pilot phase, on the basis of which the decision regarding progression to efficacy study will be made.

Intervention

MBT-PP

Tavistock Relationships' (TRs') MBT-PP is a 10-session intervention for parents which is suitable for separated parents or intact couples. The 10 sessions across 16 weeks include two assessment sessions and eight sessions that begin by introducing the skills and behaviours necessary for mentalising: the capacity to hold others in mind when emotionally aroused and to avoid a swift eruption of conflict. The subsequent sessions build on this ability to think about parents' own feelings and beliefs, those of the other parent and the needs of their children, ending with a focus on how to maintain the achievements made. Further information about MBT-PP can be found at https://tavistockrelationships.org/images/MBT_briefing.pdf.

Who: The target population for this intervention was parents of children experiencing high levels of persistent and unresolved IPC, including both separated and intact couples. The intervention was delivered to parents only – there was no direct therapeutic work with children. This evaluation focused specifically on parents of children aged 8–14 years. This age range was selected in order to fall within the range of interest to the funder (6–14 years) and also to facilitate the selection of measures which could be used across the age range. Recruitment was via referrals in three local authorities: Dorset Council; Bournemouth, Christchurch and Poole (BCP) Council; and Bristol City Council.

What: The intervention consisted of MBT-PP delivered online. After an initial assessment period, which consisted of individual sessions for each parent with the therapist, both parents took part in joint online sessions with the therapist. If only one parent was willing to take part, the therapist worked with that parent on their own. MBT-PP was delivered by MBT-PP practitioners, all of whom were qualified therapists, counsellors, family therapists, or child and adult psychotherapists, accredited and registered with their relevant professional bodies (British Association for Counselling and Psychotherapy, UK Council for Psychotherapy, Association for Family Therapy and Systemic Practice, etc.), and compliant with the requirements of these professional bodies, including ethical standards and professional supervision. Treatment fidelity was supported through fortnightly group supervision offered by MBT-PP supervisors who had received additional training. Supervisors attempted to ensure adherence to and prevent departure from the manualised intervention. This allowed TR to maintain fidelity and clinical oversight, manage risk, and develop practitioners' skills further. Supervisors' work, in turn, was overseen by monthly supervision delivered by the most experienced MBT-PP leaders.

How much: The intervention consisted of 10 sessions delivered over (approximately) 16 weeks. The first two sessions lasted up to 75 minutes, and sessions three to seven lasted 60 minutes. Sessions were usually delivered weekly or fortnightly, allowing participants some scheduling flexibility. Sessions were delivered online via secure Zoom or Teams calls, depending on client preference.

The logic model for MBT-PP can be found in Appendix A. The logic model was based on a theory of change which was developed jointly and iteratively between the deliverers and the research team during a period of co-design prior to submitting the first version of the trial protocol to the funder (and subsequently securing the funding). To generate the logic model, the requirements for delivery of the project according to the theory of change were identified and mapped onto the funder's specified progression criteria requirements.

Treatment as usual

Because of the aforementioned underdeveloped pathways for tackling IPC in LAs, it was anticipated that true TAU, that is, what would occur in the absence of this trial, would vary extremely widely. Because of the nature of family support provision by LAs, understanding the precise nature and content of all the available interventions was not possible. For example, we would not be able to appraise the intensity or content of family support worker visits to a family home in which the impact of IPC on the children may either become a focal point or be tackled much more indirectly, if at all. In this scenario, it would be very difficult to understand whether the active ingredients of MBT-PP are, in fact, unique to the intervention received in the MBT-PP arm.

Two of the participating LAs (Dorset and BCP) were already using a suite of digital resources for IPC produced by OnePlusOne (OPO), a charity that focuses on the development of healthy relationships. These resources are a low-intensity intervention, designed to be best used in a self-guided capacity spread over several weeks, rather than as pure self-help. The suite comprises three digital resources, two of which are relevant to the target group: *Arguing Better* (AB) is targeted at couples experiencing conflict, and *Getting It Right For Children* (GIRFC) is targeted at separating/separated parents experiencing conflict. The resources, widely used across England and Wales, were based on behaviour modelling training, which is distinct from and shares no overlap with the psychoanalytic underpinnings of MBT-PP.

There is no single digital programme for reducing parental conflict that is suitable for both intact and separated parents. This is because digital interventions rely heavily on scenarios and examples that need to be relatable to parents' own experiences. Conversely, live, face-to-face interventions delivered by a therapist (either in person or via video call) have the scope to be tailored to individual circumstances and idiosyncrasies while still adhering to the intended therapy. This means that while MBT-PP could be used with both intact and separated parents, digital programmes (which do not involve a live therapist) could not be so flexible.

Two of the LAs were already using the OPO programmes, hence offering the research team an opportunity to standardise, as far as possible, the content of TAU to involve known content. TAU, therefore, involved LA staff using the OPO digital resources in their work to support families that were referred to the project but had been randomised to not receive MBT-PP. LA staff used GIRFC and AB as appropriate, depending on the parental relationship status. This meant that TAU offered an appropriate intervention, targeted at IPC, but one that was far less intense than and from a different theoretical standpoint to MBT-PP. While it is likely that TAU also involved other interventions as appropriate (for example, a family needing housing advice would also have received housing advice), positioning the OPO programmes in TAU ought to have minimised variations in LA staff practices *as they relate to IPC specifically* in TAU. LA staff were not asked to withhold any specific forms of support from those allocated to the TAU arm of the trial.

Who: As mentioned above, the OPO resources were for parents of children experiencing high levels of persistent and unresolved conflict in the three specified LAs, and the intervention was delivered to parents only.

What: The TAU intervention consisted of either AB for intact couples or GIRFC for separating/separated parents. Parents were provided with login information for the relevant OPO resource and were supported in working through the online materials at their own pace. Support was provided by staff from the LA; in some cases, this was the 'gateway lead' (GL) or a member of their team, but if the family already had an

allocated family worker, then that practitioner would provide support. If only one parent was willing to take part, the practitioner supported that parent on their own, and if parents were separated, support was provided to each parent individually. Support was provided in person or remotely, depending on family preference and practitioner availability. Some families opted to work through the programme in a self-directed manner (i.e. without support).

How much: The OPO resources consist of modules which are designed to be completed over three to five sessions, taking approximately 2.5 hours to work through in total. As the resources were online, parents could work through these in their own time, with support appointments from the practitioner being arranged accordingly. Parents received up to six hours of contact time with their practitioner.

Research aims

In line with the funder's guidance on conducting pilot studies (YEF, 2022), the aims of this internal pilot study were to assess evaluation feasibility, evidence of promise and overall readiness for trial. As such, the following aims were specified in the trial protocol (Millings et al., 2024):

Evaluation feasibility

1. To assess the extent to which the referral pathways are working, i.e. whether sufficient referrals are flowing into the project and whether these referrals are meeting eligibility requirements (specified in the protocol)
2. To assess the acceptability of the referral pathways and consent and randomisation procedures to participants (indicated by dropout rates at these points)
3. To assess whether there are any signs of problematic attrition (e.g. that might indicate that the research processes or interventions are not acceptable to participants)
4. To assess how the parameter estimates used for the sample size calculation should be adjusted in light of data
5. To explore the similarity between parents' Strengths and Difficulties Questionnaire (SDQ) reports on their child(ren). This exploratory question was designed to inform the development of strategies to handle missing data, i.e. whether it would be appropriate to substitute the father's SDQ report in cases where the mother's is missing.
6. To pilot the data collection methods (including examining completion time, parent views on completing measures and missing data)
7. To pilot data linking processes

Evidence of promise

8. To seek early evidence supporting the theory of change

Overall readiness for trial will be assessed on the basis of how the above aims are met, and the red/amber/green ratings of the progression criteria below.

Progression criteria

As this is an internal pilot study, a decision regarding progression to a full efficacy study will be made on the basis of this report. Progression criteria and their associated red/amber/green thresholds were set out in accordance with the funder's guidance and outlined in the protocol for the different domains of project

implementation (Table 1), evaluation measurement (Table 2), and measurement and findings (Table 3). The red/amber/green ratings for each have been applied to the data from families recruited before 01 May 2024. It should be noted, however, that while the number of families recruited by this time was in the amber category, with 90 families randomised, the sample size at the time of writing Version 1 of this report was 110 families (159 children and young people [CYPs]), which would be in the green category. Because it was possible that attrition from the evaluation at the point of post-intervention measures (rated red) may have continued to reduce as data were still coming in, attrition figures (criterion 9) were updated on 19 September 2024.

Table 1: Project implementation progression criteria

Area	Question	Progression criteria	RAG rating
1. Fidelity	Are therapy sessions being recorded?	Percentage of therapy sessions delivered that are recorded: • Red: <50% • Amber:= 50–74% • Green: ≥75%	95.8% (227 sessions recorded of 237 sessions provided).
2. Eligibility/referral	Do enough referrals meet the eligibility criteria?	Percentage of referrals received during months 9–16 (Sept 2023 to April 2024) of the pilot that met the eligibility criteria: • Red: <50% • Amber: 50–74% • Green: ≥75%	75.3% (197 of 260 referrals).
3. Dosage	Do MBT-PP clients attend enough therapy sessions?	i) Percentage of clients who were discharged having attended six or more sessions of MBT-PP: • Red: <50% • Amber:= 50–74% • Green: ≥75%	73.5% (25 of 34 families discharged).
		ii) Percentage of clients who entered the treatment phase and were discharged having attended six or more sessions of MBT-PP: • Red: <50% • Amber:= 50–74% • Green: ≥75%	89.3% (25 of 28 families discharged who entered treatment phase).
4. Practitioner training	Have MBT-PP therapists received enough training?	Amount of the five days of post-qualification training on MBT-PP received by therapists prior to delivering MBT-PP: • Red = one or more therapists received less than three days of training • Amber = one or more therapists received only three or four days of training • Green = all therapists received the full five days of training	All MBT-PP therapists working as part of this project (n = 8) received five days of post-qualification training.

5. Supervision	Do MBT-PP therapists receive enough supervision?	Mean hours of supervision per month received by therapists delivering MBT-PP: • Red = one or more therapists received <1 hour of supervision per month • Amber = one or more therapists received >1 hour but <1.5 hours of supervision per month • Green = all therapists received ≥1.5 hours of supervision per month	All therapists received ≥1.5 hours of supervision per month between July 2023 and April 2024 (mean of 2.64 hours per month).
6. Practitioner capacity	Do MBT-PP therapists have the capacity to work with clients as intended?	Percentage of families (or individual parents if parents are attending therapy separately) that are contacted by the therapist within 2 weeks of being allocated to a therapist. • Red = <50% • Amber = 50-74% • Green = ≥75%	95.6% of families were contacted by their allocated therapist within 2 weeks.

Table 3: Evaluation measurement progression criteria

Area	Question	Progression criteria	RAG rating
7. Overall recruitment to evaluation.	Have enough families been recruited? Is the project on track to meet the recruitment needed for the efficacy study?	i) Comparison of actual vs. required recruitment. By the end of the pilot period, 140 families are expected to have been recruited and randomised. • Red = <50% of target (<69 families) • Amber = 50-74% of target (70-104 families) • Green = ≥75% of target (105 families)	90 families recruited = 64.3% of target.
		ii) % of recruited families where both parents are participating • Red = <55% of recruited families • Amber = 55-79% of recruited families • Green = ≥80% of recruited families	64.4% (58 of 90 families)
8. Attrition from MBT-PP.	Have enough families that started MBT-PP completed the treatment protocol?	Percentage of families who attended session 1 who were discharged before completing treatment protocol, i.e., before reaching the 'ending and signposting phase' in sessions 9 and 10. • Red = ≥50% • Amber = 30-50% • Green = <30%	26.5% (9 of 34 families discharged).
9. Attrition from the evaluation.	Have enough parents and CYP (who were recruited and randomised) completed the post-treatment outcome measures? Does attrition differ systematically	i) Attrition of Parent 1s (provider of primary outcome measure) across study arms: • Red: >30% • Amber: 11–30% • Green: ≤10%	4 of 37 Parent 1s who completed treatment = 10.8%.
		ii) Percentage of Parent 1s who were lost from evaluation and who were in MBT-PP condition (aiming for approx. half): • Red: <35% or >65% • Amber: 35–44% or 56–65% • Green: 45–55%	1 of 4 Parent 1s lost from evaluation = 25%.

	between study arms?	iii) Attrition of Parent 2s (who completed baseline measures) across study arms: • Red: >30% • Amber: 11–30% • Green: ≤10%	13 of 28 Parent 2s who completed treatment = 46.4%.
		iv) Percentage of Parent 2s who were lost from evaluation and who were in MBT-PP condition (aiming for approx. half): • Red: <35% or >65% • Amber: 35–44% or 56–65% • Green: 45–55%	6 of 13 Parent 2s lost from evaluation = 46.2%.
		v) Attrition of CYP (who completed baseline measures) across study arms: • Red: >30% • Amber: 11–30% • Green: ≤10%	7 of 47 CYP in families that completed treatment = 14.9%
		vi) Percentage of CYP who were lost from evaluation and who were in MBT-PP condition (aiming for approx. half): • Red: <35% or >65% • Amber: 35–44% or 56–65% • Green: 45–55%	3 of 7 CYP lost from evaluation = 42.8%.

Note: RAG = red/amber/green; MBT-PP = Mentalization Based Therapy for Parenting under Pressure; CYP = children and young people

Table 3: Measurement and findings progression criteria

Area	Question	Progression criteria	RAG rating
10. Feasibility of randomisation	Did randomisation work? Were there any problems?	Percentage of participants allocated to MBT-PP condition: • Red: <35% or >65% • Amber: 35–44% or 56–65% • Green: 45–55%	50%.
		Problems with the randomisation process as reported by SHU, TR and LAs: • Red: significant problems, major changes to processes needed • Amber: minor problems, refinements to processes needed • Green: no problems	No problems
11. Data quality	Are the baseline and post-treatment primary outcome measure data of high quality?	Percentage of missing data within the 25 SDQ items completed by parents at baseline and post-treatment: • Red: >30% • Amber: 10–30% • Green: ≤10%	Baseline: 0.04% (two missing items out of 5,475 SDQ items). Post-treatment: 0% (0 missing items out of 1,150 items).

12. Effective use of core measures	Is there an effective mechanism in place to collect the outcome measures?	Mechanisms in place to collect outcome measures from parents and CYP: • Red: there is no mechanism in place to support data collection in each LA • Amber: there is not a GL currently working in each LA, but alternative arrangements to support data collection have been or are being made where necessary • Green: a GL is in post in each LA to liaise with the evaluator for the completion of baseline, post-treatment and follow-up outcome measures	A GL is in post in each LA to liaise with the evaluator for the completion of baseline, post-treatment and follow-up outcome measures.

Note: RAG = red/amber/green; MBT-PP = Mentalization Based Therapy for Parenting under Pressure; CYP = children and young people; GL = Gateway Lead; SDQ = Strengths and Difficulties Questionnaire; SHU = Sheffield Halam University; TR = Tavistock Relationships; LA = local authority

Overall, the pilot study shows good readiness for trial, with 12 criteria rated green, six rated amber (one of which cannot be green due to the small, odd numbers involved and because green requires a closer-to-50% proportion than possible) and two rated red. The amber criteria were related to recruitment and attrition, and the red criteria were related to attrition. We have identified additional strategies to address attrition so that we can improve these criteria in the main trial (see the section on evaluation feasibility).

Ethical review

Ethical approval was obtained from Sheffield Hallam University (SHU) University Research Ethics Committee, which has established research ethics policies and procedures aligned with legal requirements and research societies' standards of good practice (https://www.shu.ac.uk/research/quality/ethics-and-integrity).

The research protocol, together with a summary of the study methods and procedures and all participant-facing documentation (including participant information sheets, consent forms and measures), were subjected to review by three trained anonymous reviewers. Ethical approval was granted (reference number ER50582599).

Participants reviewed the information sheets (Appendices B and C) prior to providing their consent to participate. Informed consent was provided by the parents for their own participation. Informed consent was provided by the parents for the participation of their children, who also assented to participate.

No ethical problems were encountered in the pilot phase.

The trial was registered: ISRCTN10266960.

Data protection

SHU undertakes research as part of its function for the community under its legal status. Data protection laws allow us to use personal data for research with appropriate safeguards in place under the legal basis of public tasks that are in the public interest. Information about the University's legal status, constitution and public tasks can be found here: https://www.shu.ac.uk/about-us/governanceand-strategy/governance/legal-status-and-constitution.

In alignment with YEF's race equity commitment, we aim to ensure that race, ethnicity and inclusion are a key focus throughout this project. To this end, we collect information on racial and ethnic origin, which falls under the category of special category data under GDPR UK. As mentioned above, our legal basis for processing these data is that it is necessary for reasons of substantial public interest. Participants are given the option to not provide these data should they so wish.

We inform potential participants about the information we wish to collect from them and how we will use it. We seek their consent for the collection and use of their data. The Research Ethics Committee agreed that our consent procedure is appropriate and ensures that participants' rights are protected. Full details are provided to participants (as well as their parents/legal guardians in the case of CYP) in an information sheet.

Research at the University is governed by policies and procedures, and all research undergoes ethical scrutiny to ensure that it is conducted in such a way as to protect participants' interests and is of a high standard (https://www.shu.ac.uk/research/ethics-integrity-and-practice).

We only collect information that is essential for the purpose of the research. Research data are treated as confidential and are identifiable only via ID numbers. The key that links ID numbers with individuals is stored securely and separately from the research data and is accessible only by authorised individuals in the research team for the current trial.

The privacy of personal data is paramount, and no personal data will be disclosed unless there is a justified purpose for doing so. The University NEVER sells personal data to third parties. Data may be shared with:

- The immediate project team, which is authorised to work on the project and access the information. This may include staff at SHU or collaborators at other organisations authorised to work on the project. This will be clearly identified in the information sheet. Our research may be audited, and access to the data may be required. The University puts in place safeguards to ensure that audits are conducted in a secure and confidential manner.
- In the case of complaints about a research project, the Head of Research Ethics may require access to the data as part of our Research Misconduct Procedure.

The University takes a robust approach to protecting the information it holds, with dedicated storage areas for research data with controlled access. For particularly sensitive projects, the University puts additional layers of security into place. The University has a high level of data security and follows the National Cyber Security Centre's '10 Steps to Cyber Security' framework to structure security planning and operations. Through information strategy, policy and process, the University is aligning with the ISO27001 standard. Alongside these technical measures, there are comprehensive and effective policies and processes in place to ensure that users and administrators of University information are aware of their obligations and responsibilities for the data to which they have access. By default, people are only granted access to the information they require to perform their duties. Training is provided to new staff joining the University, and existing staff have training and expert advice available if needed.

Research data will be prepared for archiving in the YEF evaluation data archive at the end of the project. Two data sets will be created. One file will contain just the CYP's identifying data. This will be submitted to the Department for Education (DfE), where personal data will be removed. This will be replaced with the DfE's pupil matching reference numbers (PMRs) and then submitted to the Office for National Statistics

(ONS) for storage in the YEF archive. The second file will contain the evaluation data. This will be submitted directly to the ONS and stored in the YEF archive. It will contain unique identifiers that allow it to be connected with the DfE PMRs. Both data sets will be held securely in the YEF archive by the ONS.

Future approved researchers will be able to access the data for approved research projects and link it with the DfE's National Pupil Database (for data on school exclusions and academic performance) and the Police National Computer (for criminal justice information).

No data protection problems were encountered in the pilot phase.

Project team/stakeholders

Project team roles and responsibilities (all affiliated to TR):

Andrew Balfour – TR CEO

Sarah Ingram – associate director, responsible for partnerships and delivery, and TR project leader and YEF contact, responsible for project management, good governance and delivery accountability, working with SHU colleagues to offer feedback on progress to LAs

Maria Franchini – MBT lead, responsible for MBT-PP delivery, fidelity, supervision and further training as required

Evaluation team roles and responsibilities (all affiliated to SHU):

Prof. Abigail Millings – evaluation lead, responsible for oversight of every aspect of the trial

Sean Demack and Prof. John Reidy – evaluation statisticians. Sean will lead on the primary and secondary outcomes analyses, and John will lead on the mechanisms of action analyses.

Dr Charlotte Coleman – advisor for IPE with a focus on working with CYP

Prof. Madelynne Arden – advisor for treatment fidelity, compliance and contamination

Dr Elaine Clarke – trial manager, responsible for data collection infrastructures, data management and interviews

Dr Laurynas Rutkauskas – researcher, responsible for allocations, data quality, data management and interviews

Piyali Misquitta – PhD candidate, responsible for allocations and interviews

Anna Stevens – advisor for data amalgamation and cleaning processes

Methods

Trial design

This internal pilot study was the first phase of an intended pragmatic two-arm cluster RCT in which the unit of randomisation is the family. The trial was designed by the evaluators with input from the deliverers to ensure feasibility. The resultant protocol was registered: ISRCTN10266960.

Implementation

The developer funded the LAs for their staff time on the project (2 x 0.6 FTE posts per LA) and their licenses to use OPO's digital materials for TAU from the developer's YEF project delivery budget. As such, the LAs were accountable to the developer. The evaluator funded the participant incentives from the evaluator's YEF evaluation budget. The evaluator trained and instructed the LA postholders on how to fulfil the LA functions for the trial, including referrals, consent and recruitment, and screening for eligibility. Both the developer and the evaluator worked with the LA postholders and their managers to design referral pathways that worked in the local contexts while adhering to the requirements of the trial. Both the developer and the evaluator met with the LA postholders weekly to provide ongoing support and assistance. LA postholders also supported each other in these sessions. Issues raised tended to be around the application of the eligibility criteria in complex circumstances, and these were resolved via discussion between all parties. The best-performing LA (in terms of recruitment numbers) shared its implementation methods with the remaining two LAs to share learning and promote best practices around managing referrals to facilitate as high a conversion rate into research participation as possible.

The trial design remained per protocol throughout the pilot, with some minor exceptions. The eligibility criteria were relaxed to permit i) families in which only one parent wanted to engage initially and ii) families with CYP who would turn eight during the intervention phase or CYP who were still 14 at the time of recruitment. The rationale for these changes was to facilitate recruitment.

While the theory of change for MBT-PP indicates that the intervention is maximally effective when both parents engage, GLs found the requirement for both parents to join the project together to be a significant barrier to recruitment. As such, it was agreed with the funder that GLs would attempt to engage both parents in the trial, but should one parent be willing to engage while the other was unwilling, the family would be accepted into the trial and randomised to one of the two conditions, and the engaging parent would begin their intervention. The GL would contact the other parent one or two more times, using multiple methods where possible (phone, email, text), to invite them again to participate in the trial. Parents joining the trial at this point would be allocated to the same condition as their co-parent, and in the case of families in the MBT-PP arm, they would be invited to join their co-parent for the remaining therapy sessions.

It is expected that MBT-PP will still have a beneficial impact on parents and families even if only one parent engages with MBT-PP due to the potential for an increase in one parent's mentalising ability to interrupt IPC. If only one parent engages with MBT-PP, there will still be less conflict perceived by the child due to not having two parents who habitually escalate disagreements into intense and poorly resolved conflicts. Although there might be some dilution of the effect due to working with one parent instead of both, the quantity and quality of IPC is still likely to be impacted for the better. Indeed, evidence from the RPC programme (Brewin and Garlick, 2023), in which MBT-PP was one of several interventions utilised (data are

not split by intervention, but MBT-PP was by far the largest subgroup), suggests that the differences in effect size when one or both parents engage in an intervention are quite small on relationship measures (absolute differences ranged from 0.04 to 0.25 for separated parents and from 0.0 to 0.13 for intact couples) and extremely small on most SDQ subscales (absolute differences ranged from 0.01 to 0.23 [these data were not split by relationship status]). Regarding the SDQ total difficulties scale, which is the primary outcome variable in the current trial, the RPC programme found an effect size of 0.41 when one parent engaged and 0.51 when both parents engaged (so an absolute difference of 0.10).

Allocation was achieved using a minimisation protocol in MinimPy (an open-source computer program; Saghaei and Saghaei, 2011) to achieve balance across the following demographic variables: age group (all CYP aged 8–11 / all CYP aged 12–14 / CYP aged 8–11 and 12–14), minority ethnic group status of one or both parents (yes/no) and relationship status (separated/intact). Where a family had only one parent engaging, this meant that we were unable to collect ethnicity data on the non-engaging parent; hence, the family was entered into MinimPy for allocation based on the ethnicity of the engaging parent only. These variables were selected for minimisation as it was deemed possible that the interventions may be received differently and, hence, be differentially effective across these groupings. It was, therefore, important to achieve sample balance to retain as much control of extraneous variables as possible.

Participant selection

Participants were parents experiencing high levels of IPC, with one or more CYP aged 8–14. Participants were either referred to the project by LA staff or self-referred. Referring LA staff included family hub staff, targeted family support staff, first response staff and social workers. Self-referrals were generated by raising awareness about the project to schools, who in turn shared information with parents.

The eligibility criteria for referral were:

- Parents who were experiencing high-intensity, frequent and unresolved IPC and had at least one child aged between 8 and 14 years
- Parents did not have to live together or be in a current relationship but had to be willing to think about how they could improve their relationship with their co-parent.
- Parents needed to consent to the referral being made and have an understanding that the work would focus on the quality of the relationship with their co-parent.
- Parents needed to understand the following:
    - The project was a research project.
    - They would be randomly allocated to receive either support from the LA, which involved using an online therapeutic resource, or the MBT-PP intervention, which was delivered online.
    - They would need to fill in questionnaires before the start of their intervention, at the end, and three months after they finished.
    - They would receive vouchers to thank them for their time.
    - They would need to be willing and able to attend either 10 sessions of MBT-PP therapy or engage with the digital resources as directed by LA staff.

Referrals were received by the GL (the local postholder) from other professionals in the LA. The GL reviewed the referral and contacted the family. In the case of self-referrals only, the GL administered the four-item

version of the Dyadic Adjustment Scale (DAS-4) to assess for relationship distress. The GL also screened for risk against the exclusion criteria and checked that the inclusion criteria were met.

The inclusion criteria were as follows:

- Parents needed to have at least one child aged 8–14 (the criterion was amended to include CYP turning eight during the intervention phase, meaning not more than 16 weeks younger than eight, and CYP who were still 14 at the time of recruitment, even if they turned 15 during the research processes).
- Parents must either have been referred for support with IPC by an LA staff member or have been classified as a distressed couple on the DAS-4, with one or both parents scoring <13.

The exclusion criteria were as follows:

- Issues with substance or alcohol misuse in either parent in the previous 12 months.
- Significant mental health diagnosis which was not well managed at the time of referral: guidance was agreed between the evaluators and TR regarding how the GLs should appraise 'well managed' (see Appendix B).
- Current (at the time of referral) domestic abuse or violence (historic issues of domestic abuse/intimate partner violence needed to be detailed in the referral).
- Current (at the time of referral) engagement in court proceedings (e.g. care proceedings or private family law proceedings).

These exclusion criteria were developed in the co-design phase between the developer and the evaluators in order to ensure the safety of the participants. MBT-PP is not appropriate for abusive relationships due to the potential for the therapy sessions to provide an additional forum for the abuse or for relationships with ongoing litigation due to the potential for individuals to take and use therapy session content out of context.

Subject to the inclusion and exclusion criteria, the GL then provided participant information sheets, administered the consent procedure and ensured completion of baseline measures. This was done either in face-to-face meetings or via telephone or electronic means. An overview of the trial processes is provided in Appendix C. Participant information sheets for parents and CYP are shown in Appendices D and E, respectively.

Data collection

Pilot trial data

All consent information and quantitative outcome measure data were provided to SHU via the Qualtrics survey platform. Eligible families were supported by the GLs in completing the consent forms. This was either done in person or remotely, with either the participant or the GL entering data into the online consent form. The CYP consent form (completed by a parent) also included CYP demographics. Each family member who was participating in the trial then completed the baseline survey. Parents either did this independently or with support from the GLs, either in person or remotely. The majority of the parent surveys were filled in using the online survey platform; however, occasionally, paper copies were used either to overcome technical difficulties (e.g. no internet connection) or because of the need to use translated materials. CYP baseline measures were completed in person, in a location where the parent(s) were not present, such as

at the child's school or a family centre. If in-person surveys were filled in on paper, data were later inputted into Qualtrics by the GLs. Post-treatment and follow-up surveys were filled in in the same way, with parents either doing the online surveys independently or with support and CYP being supported in person by the GLs.

Outcome measures

Outcome measures were selected to assess the constructs identified in the theory of change, the precise links between which are further delineated in the logic model (Appendix A). Certain features of the trial design meant that there was not a wide range of measures to choose from for many of our constructs. These included the age range of the CYP targeted: 8–14 years is a wide age range and spans multiple developmental stages. Another feature that limited the measure options was the fact that families could be intact or separated. Measures of IPC are typically intended for only one of these groups and are not appropriate for both. Finally, some of our constructs were only recently identified in the literature or simply have not received similar levels of research attention as others, which means that only one measure exists (to the very best of our knowledge). Despite these challenges, we selected a set of measures that were appropriate for the target population and demonstrated relatively good psychometric properties in the pilot data. Cronbach's alphas for each scale using the pilot data are presented below.

All measures were conducted at baseline and post-intervention and will also be conducted at three month follow-up.

Primary outcome

The Strengths and Difficulties Questionnaire

The SDQ (Goodman, 1997) is being used by the YEF across its projects to create consistency and comparability between different evaluations. Further information about the SDQ is available here: https://www.sdqinfo.org/.

Parents completed the SDQ. The total difficulties score at post-intervention was the primary outcome variable. The SDQ is a brief questionnaire measuring behaviours, emotions and relationships in 4–17-year-olds. It contains 25 items within five subscales: two subscales measure externalising problems (conduct problems and hyperactivity/inattention), two measure internalising problems (emotional symptoms and peer problems) and one measures prosocial behaviour. Participants were asked to rate their children's behaviour on a series of statements from 'Not true' (0) to 'Certainly true' (2), e.g. 'Often has temper tantrums or hot tempers'. Parents rated their children's behaviour over the last six months at baseline and over the last month at post-intervention. The total difficulties score is the sum of four of these subscales, excluding prosocial behaviour, with higher scores indicating greater difficulties. The parent report questionnaire was used in this evaluation. We asked both parents to complete the SDQ for each of their children. We took mothers' SDQ scores as the primary outcome variable and treated fathers' SDQ scores as a secondary outcome.

Internal reliability of the SDQ subscales ranged from questionable ($0.60 \leq \alpha < 0.70$) to good ($0.70 \leq \alpha < 0.80$), with Cronbach's alphas of 0.72 from the mothers' data and 0.68 from the fathers' data for the conduct problems subscale, 0.81 (mothers) and 0.76 (fathers) for the hyperactivity/inattention subscale, 0.75 (mothers) and 0.77 (fathers) for the emotional symptoms subscale, and 0.66 (mothers) and 0.70 (fathers)

for the peer problems subscale. For the total difficulties scale, Cronbach's alphas were 0.66 for mothers and 0.68 for fathers.

Secondary outcomes

The Stirling Children's Well-being Scale (SCWBS; Liddle and Carter, 2015) was used to measure children's well-being. It is validated for use in children aged 8–15 years. It consists of 12 items measuring children's emotional and psychological well-being over the previous two weeks and three items to assess socially desirable responding. Participants were asked to rate their agreement with statements on a 5-point scale, from 'Never' (1) to 'All of the time' (5). All items are positively worded, e.g. 'I've been in a good mood'. Item scores are summed to produce a total, with higher scores indicating higher well-being. Baseline internal reliability was good, with a Cronbach's alpha of 0.86.

The O'Leary–Porter Scale (OPS; Porter and O'Leary, 1980) was used to assess parent-reported IPC. The OPS is a 10-item measure designed to assess overt hostility in intact couples, but it has also been used with separated couples (e.g. Owen and Rhoades, 2012; Shifflett and Cummings, 1999). Participants were asked to rate the frequency of overt hostility (such as quarrels, sarcasm, physical abuse) observed by their children on a 0–4 scale from 'Very often' (0) to 'Never' (4), e.g. 'How often do arguments between you and your child's other parent take place in front of this child?' The single negatively worded item was reverse coded, and item scores were summed to produce a total score. Lower scores indicate greater hostility in the relationship. We gained the author's permission to adapt this scale to suit a modern British audience. Internal consistency at baseline was acceptable to good, with Cronbach's alphas of 0.79 from the mothers' data and 0.82 from the fathers' data.

The Perceptions of Interparental Conflict-Intensity/Frequency Scale (PIC-I/F; Kline, Wood and Moore, 2003 is a 13-item scale measuring children's views of aspects of relationship conflict. Participants were asked to indicate how true statements were for the parents' relationships on a 6-point scale from 'Definitely false' (1) to 'Definitely true' (6). This measure is a short form of the 48-item Children's Perception of Interparental Conflict Scale (Grych, Seid and Fincham, 1992), which was developed for 9–17-year-olds. Item scores were summed, with higher scores indicating higher perceptions of IPC. The PIC-I/F had good internal reliability at baseline, with a Cronbach's alpha of 0.85.

The Dimensions of Anger Reactions–Revised (DAR-R; Nederlof et al., 2009) was used to measure the parents' anger. The seven-item scale assesses anger responses and functional impairment. Participants were asked to rate how much the statements had applied to them over the previous four weeks from 'Not at all' (0) to 'Very much' (4), e.g. 'I often find myself getting angry at people or situations'. Scores were summed to produce a total score, with higher scores indicating higher levels of anger. Internal reliability was good at baseline, with a Cronbach's alpha of 0.83.

The Parental Reflective Function Questionnaire (PRFQ; Luyten et al., 2017) was used to measure parent mentalising ability. The PRFQ is an 18-item measure consisting of three subscales: pre-mentalising modes; certainty about mental states; and interest and curiosity in mental states. Participants were asked to rate their agreement with statements on a 7-point scale from 'Strongly disagree' (1) to 'Strongly agree' (7), e.g. 'I try to see situations through the eyes of my child'. Negatively worded items were reverse coded, and mean scores were calculated to create subscale scores. Higher scores indicate higher parental mentalising ability. Baseline internal reliability of the pre-mentalising modes subscale from the mothers' data was poor, with a Cronbach's alpha of 0.56, but it was acceptable from the fathers' data, with a Cronbach's alpha of 0.73. The

Cronbach's alphas for the certainty about mental states subscale were 0.78 (mothers) and 0.65 (fathers), and for the interest and curiosity about mental states subscale were 0.72 (mothers) and 0.81 (fathers).

The Parenting Scale Short Form (PS-8; Kliem et al., 2019) was used to measure parenting style. The PS-8 uses eight items from the Parenting Scale (Arnold et al., 1993), which assesses parenting behaviour in response to problematic child behaviour over the previous two months. Participants were asked to rate their agreement with statements on a 7-point scale between two poles, representing effective and ineffective parenting strategies. Mean scores were calculated, with higher scores indicating more dysfunctional parenting. Baseline internal consistency was acceptable to good, with Cronbach's alphas of 0.79 from the mothers' data and 0.80 from the fathers' data.

The Emotional Adaptation to Relationship Dissolution Assessment (EARDA; Millings et al., 2020) is a 10-item scale developed in the UK and validated in samples of separated parents. The EARDA has excellent convergent, discriminant, concurrent criterion-related and incremental validity; correlates with co-parenting communication; mediates between separation characteristics and conflict; and, in a small sample, has been found to align with the professional opinions of mediators regarding parents' ability to communicate without arguing (Millings et al., 2020). Because the EARDA's focus is an adaptation to the dissolution of the relationship, it was used only for parents who were separated. Participants were asked to rate the accuracy of a series of statements from 'Does not describe my feelings at all' (0) to 'Describes my feelings exactly' (5), e.g. 'I feel a failure that my relationship broke down'. Negatively worded items were reverse coded, and mean scores were calculated, with higher scores indicating higher emotional adaptation to the relationship dissolution. The Cronbach's alpha was acceptable at baseline, at 0.71.

Socio-demographic information

In addition to the validated outcome measures listed above, parents were asked to report socio-demographic information at baseline, e.g. age, sex, gender, ethnicity, LGBTQ+ status, education level, marital status, employment status and postcode. We also asked the participants to report their children's ethnicity and sex, whether their children had any special educational needs or disabilities (SEND), and whether they were looked-after children/previously looked-after children. We collected children's dates of birth and home addresses to allow the children to be matched at the end of the project to the records held in the National Pupil Database by the DfE.

Mentalization Based Therapy for Parenting under Pressure administrative data

For participants allocated to the MBT-PP condition, we collected data from TR about the date of first contact with each family, how many therapy sessions were attended per family, how many therapy sessions were delivered and recorded, and how many hours of supervision therapists received.

Interviews with Gateway Leads

All GLs (n = 3) were invited for an interview to discuss their experiences of working on the trial so far. Interviews were semi-structured, using a topic guide developed to meet the needs of the feasibility-related research aims. Interviews were conducted at times convenient to the participants after they had provided informed consent. Interviews were conducted by a member of the research team who had not been working directly with the GLs and lasted 40 minutes to an hour each. Interviews were fully transcribed prior to analysis.

Qualitative analysis

Transcripts were analysed using a realist, pragmatic, directed, qualitative content analysis approach (Assarroudi et al., 2018), focusing on research questions (RQs) 1, 2 and 5 (Table 4). The approach taken was pragmatic, as the analysis was guided by the need to surface any insights that would help identify problems, suggest solutions or generally streamline trial processes, with a view to improving them for the efficacy phase.

Trial feasibility data

Multiple data sources and methods were used to address the research questions (RQs) regarding the feasibility of the efficacy study (Table 4).

Table 4. Methods overview

Research questions addressed	Participants/data sources	Data analysis method	Logic model relevance
1. To what extent are the referral pathways working?	Referral and recruitment rates and proportion of referrals meeting eligibility criteria Interviews with GLs	Descriptive graphs Thematic analysis	Progress criteria 2 and 7, specified in the logic model (Appendix A)
2. To what extent are the referral pathways, consent and randomisation procedures acceptable to participants?	Dropout rates and location in the 'pipeline' Interviews with GLs	CONSORT flow diagram Qualitative content analysis	Progress criteria 2 and 7, specified in the logic model (Appendix A)
3. Are there any signs of problematic attrition?	Dropout rates and position in the 'pipeline'	CONSORT flow diagram	Progress criteria 2 and 7, specified in the logic model (Appendix A)
4. How should the estimates used for the sample size calculation be adjusted in light of the data?	Demographic and outcome data from trial variables	Descriptive statistics Tests of association Intra-class correlations at the CYP and family levels	n/a
5. Do the data collection methods work?	Data quality Scale reliabilities Measure completion time Interviews with GLs Piloting of fidelity checklist	Percentage of SDQ data missing Cronbach's alphas Mean response times on Qualtrics Thematic analysis Description of process	n/a
6. Do data linking processes work?	Datasets of quantitative data collected	Description of process	n/a
7. Is there any early evidence supporting the theory of change?	Outcome data from trial variables	Descriptive statistics by trial arm	The direction of change anticipated on each variable is specified in the logic model (Appendix A)
8. What is the relationship between parents' SDQ reports?	Outcome data from trial variables	Tests of association	n/a

Note: GL = Gateway Lead; SDQ = Strengths and Difficulties Questionnaire; CYP = children and young people

Randomisation

Randomisation was conducted on all consenting referrals as they occurred. Because recruitment was rolling, classic random stratification was not feasible. We, therefore, used minimisation (Scott et al., 2002; Altman and Bland, 2005), whereby allocation was initially random and then systematic to minimise differences between groups across a few specified strata within LAs, specifically, child age range, parent minority ethnic group status and relationship status (intact or separated). In the context of rolling recruitment, a minimisation approach best ensures that the MBT-PP and TAU samples are comparable in terms of the specified strata. The LA of the family was not used as a minimisation variable, but for families randomised before 1 May (n = 90), there was no significant association between the family's LA and the allocated condition: $\chi^2(2) = 0.87$, $p = 0.65$.

Participants were enrolled into the trial by the GLs in each of the three LAs. The random allocation was generated by research staff at SHU using MinimPy software. The research staff then informed the relevant GL of the condition to which each family had been allocated.

The unit of randomisation was the family. Randomisation of each family only took place after consent forms and baseline measures had been obtained for all family members who were participating in the trial and after demographic data had been obtained for each relevant child (in the age range and affected by the IPC), even if the child were not completing outcome measures themselves. From these data, each family was classified on each of the three strata. For child age, families fell into one of three categories: i) children aged 8–11 years, ii) children aged 12–14 years or iii) children aged 8–11 years and 12–14 years (i.e. families in which there was a sibling in each age bracket). For parent minority ethnic status, families were categorised as either i) yes, where at least one parent was from an ethnic minority background, or ii) no, where neither parent was from an ethnic minority background. For relationship status, families were categorised as either i) intact or ii) separated based on the parents' self-reports.

Once each family had been classified on the three strata, these data were input into the MinimPy software, which allocated that family to either MBT-PP or TAU on a 1:1 basis. When more than one family was input into MinimPy on the same day, the order in which the families were entered into the programme by the research team was decided on a random basis (with a coin toss if allocating two families or using an online random number generator if allocating more than two families on one day). The allocated condition for each family was recorded, and the GL working with that family was informed of the outcome of the allocation so that they could relay this to the family. TR was given only the parent contact details for families that had been allocated to MBT-PP. There was no possibility that either the families themselves or the GLs could predict which condition a family would be allocated to, as no information was shared between SHU and the LAs about the ongoing allocations of all families in the trial, only about individual families on a rolling basis. The GLs were blind to the allocation of families outside their LAs.

Two scenarios emerged during the pilot which were not anticipated in advance, so the following strategies were developed when needed. First, in cases in which there was a discrepancy between parents about whether their relationship was intact or they were separated, families were classed as separated. Second, in families in which only one parent participated in the trial, families were categorised based on the sole parent about whom we had data.

Timeline

Table 5 presents the timeline of key activities for the pilot phase as they happened.

Table 5: Timeline

Date	Activity
07/02/2023	GL in Bristol in post – Bristol LA. Admin support provided since January 2024.
01/03/2023	Intended start to recruitment/referrals
21/04/2023	Start of recruitment – GLs (where appointed) begin to advertise the project as accepting referrals.
07/06/2023	GL and admin in Dorset in post (but admin left post after one month). Replacement admin support in post from October 2023.
14/06/2023	GL in BCP in post. Admin support post not provided.
21/06/2023	Start of data collection (first family that provided data) – GLs
29/06/2023	First family randomised – SHU research team
29/06/2023	Start of intervention delivery – TR (and LAs for TAU).
03/10/2023	Admin support in post in Dorset.
13/10/2023	Proposed changes to protocol and pilot period – SHU research team and TR
24/10/2023	Request to extend pilot period, due to end in December 2023, to April 2024 in order to account for delays in LAs recruiting GLs and to include one-parent families in order to remove barriers to recruitment, approved – YEF.
09/01/2024	End of recruitment for pre–post data collection within pilot
22/01/2024	Admin support in Bristol in post.
30/04/2024	End of data collection for pilot.
01/05/2024	Data cleaning and wrangling. Creation of CYP-level dataset from parent and CYP survey data across time points. Creation of parent-specific dataset – SHU research team
01/05/2024	Collection of MBT-PP administrative data from TR – SHU research team and TR
20/05/2024	Data analysis by SHU research team for pilot report

Note: GL = Gateway Lead; LA = local authority; SDQ = Strengths and Difficulties Questionnaire; CYP = children and young people; BCP = Bournemouth, Christchurch and Poole; TR = Tavistock Relationships; TAU = treatment as usual; YEF = Youth Endowment Fund; MBT-PP = mentalization-based therapy for parenting under pressure

Quantitative analysis

Given that it is not appropriate to conduct tests of difference between trial arms for a pilot study (Eldridge et al., 2016), the analysis utilised descriptive statistics for all sample characteristics and quantitative variables. Tests of association were used to address RQs 4 and 8 (Table 4).

Findings

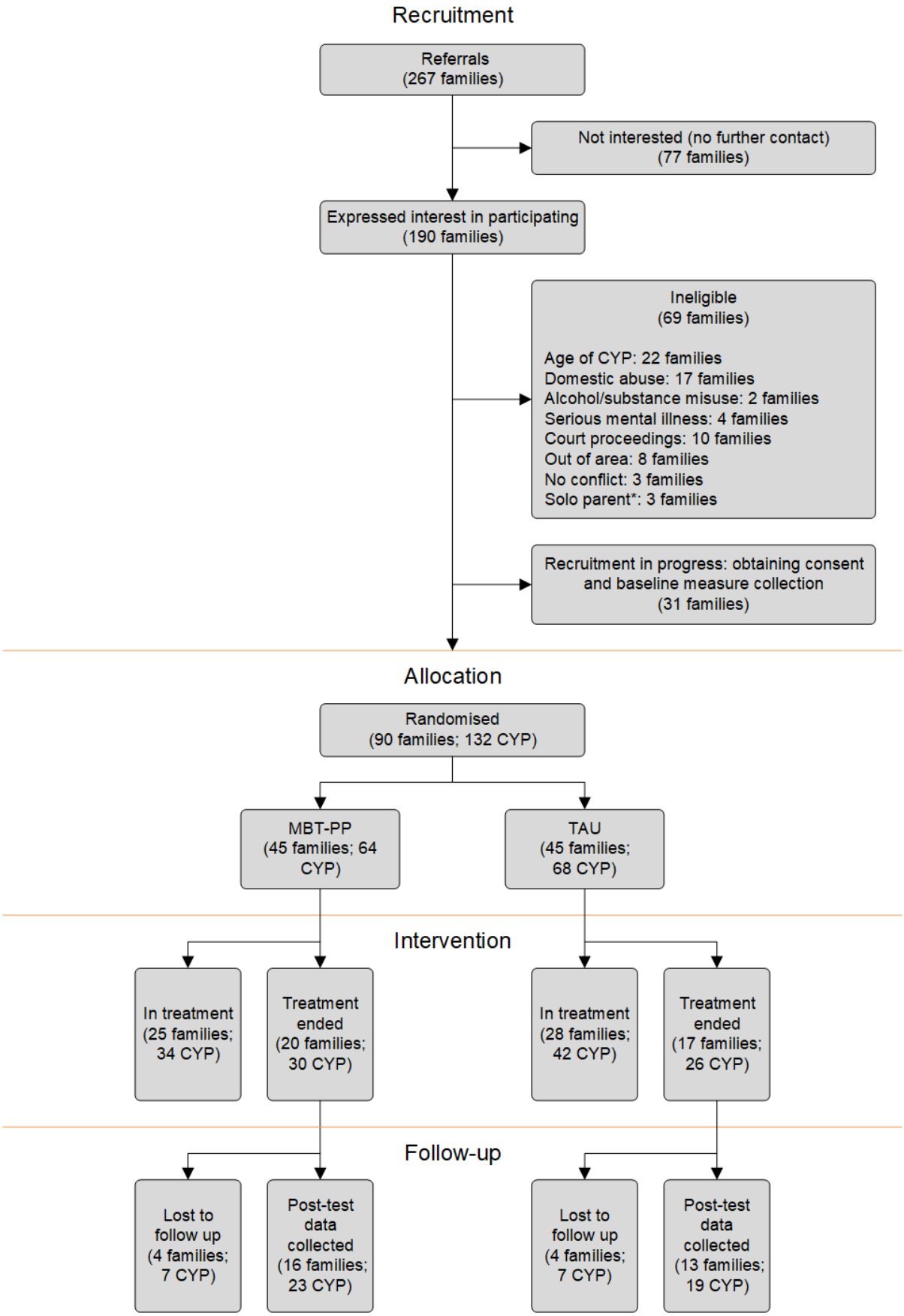
Participants

The numbers of CYP in the flow diagram (Figure 1) refer to the number of CYP (aged 8–14) in participating families about whom we collected data, regardless of whether or not the CYP themselves completed secondary outcome measures. Numbers for post-test data collected refer to the primary outcome measure (i.e. mothers' SDQ scores).

Ethnicity data for parents and CYP by LA are provided in Table 6.

The three locations participating in this evaluation all have a majority White British population, and this is reflected in our sample characteristics.

Figure 1: Participant flow diagram



*Prior to the inclusion of one-parent families.

Table 6: Ethnicity data by local authority

Ethnic background	Bristol		Dorset		BCP	
	Parents	CYP	Parents	CYP	Parents	CYP
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
White British	33 (73.3%)	27 (81.8%)	57 (96.6%)	54 (96.4%)	30 (66.7%)	29 (67.4%)
White Other	4 (8.9%)	1 (3.0%)	1 (1.7%)	2 (3.6%)	11 (24.4%)	11 (25.6%)
Mixed/Multiple ethnic groups	3 (6.7%)	4 (12.1%)	0 (0%)	0 (0%)	2 (4.4%)	2 (4.7%)
Asian/Asian British	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (4.4%)	1 (2.3%)
Black/Black British	1 (2.2%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Other ethnic group	4 (8.9%)	1 (3.0%)	1 (1.7%)	0 (0%)	0 (0%)	0 (0%)
Total	45 (100%)	33 (100%)	59 (100%)	56 (100%)	45 (100%)	43 (100%)

Note: BCP = Bournemouth, Christchurch and Poole

Demographic data for parents at baseline by condition is presented in Table 7. A little over half of the parents (55.7%) in our sample were in intact families at baseline (i.e. the parents experiencing conflict were still in a relationship), and 57.7% of parents were female. The most common marital status of parents was married or in a domestic partnership, and the large majority (96.6%) described their sexual orientation as heterosexual. Families contained a mean of 1.47 CYP.

Table 7: Parent baseline sociodemographic characteristics (n = 149)

Characteristic	MBT-PP group (n = 73)	TAU group (n = 76)	All parents
Gender, n (%)			
Male	30 (41.1%)	32 (42.1%)	62 (41.6%)
Female	42 (57.5%)	44 (57.9%)	86 (57.7%)
Non-binary	1 (1.4%)	0 (0.0%)	1 (0.7%)
Ethnicity, n (%)			
White British	58 (79.4%)	62 (81.6%)	120 (80.5%)
White Irish	0 (0.0%)	1 (1.3%)	1 (0.7%)
Any other White background	8 (11.0%)	7 (9.2%)	15 (10.1%)
Mixed or multiple ethnic groups	4 (5.5%)	1 (1.3%)	5 (3.3%)
Asian or Asian British	0 (0.0%)	2 (2.6%)	2 (1.3%)
Black or Black British, Caribbean or African	1 (1.4%)	0 (0.0%)	1 (0.7%)
Other ethnic group	2 (2.7%)	3 (3.9%)	5 (3.4%)
Sexual orientation, n (%)			
Heterosexual	70 (95.9%)	74 (97.4%)	144 (96.6%)

Characteristic	MBT-PP group (n = 73)	TAU group (n = 76)	All parents
Bisexual	2 (2.7%)	2 (2.6%)	4 (2.7%)
Other sexual orientation	1 (1.4%)	0 (0.0%)	1 (0.7%)
Relationship status, n (%)			
Intact	41 (56.1%)	42 (55.3%)	83 (55.7%)
Separated	32 (43.8%)	34 (44.7%)	66 (44.3%)
Marital status, n (%)			
Single	11 (15.1%)	11 (14.5%)	22 (14.8%)
Married/domestic partnership	46 (63.0%)	45 (59.2%)	91 (61.1%)
Widowed	0 (0.0%)	1 (1.3%)	1 (0.7%)
Divorced	7 (9.6%)	7 (9.2%)	14 (9.4%)
Separated	9 (12.3%)	12 (15.8%)	21 (14.1%)
Employment status, n (%)			
Employed full time	31 (42.5%)	27 (35.5%)	58 (38.9%)
Employed part-time	19 (26.0%)	21 (27.6%)	40 (26.8%)
Self-employed	12 (16.4%)	11 (14.5%)	23 (15.4%)
Unemployed	4 (5.5%)	8 (10.5%)	12 (8.1%)
Unable to work	7 (9.6%)	4 (5.3%)	11 (7.4%)
Other	0 (0.0%)	5 (6.6%)	5 (3.4%)
Education level, n (%)			
No qualifications	0 (0.0%)	2 (2.6%)	2 (1.3%)
GCSE or equivalent	13 (17.8%)	10 (13.2%)	23 (15.4%)
A and AS level or equivalent	10 (13.7%)	11 (14.5%)	21 (14.1%)
Apprenticeships	1 (1.4%)	0 (0.0%)	1 (0.7%)
NVQ or equivalent	18 (24.7%)	23 (30.3%)	41 (27.5%)
Degree level or higher	29 (39.7%)	29 (38.2%)	58 (38.9%)
Other	2 (2.7%)	1 (1.3%)	3 (2.0%)
Local authority, n (%)			
Bristol	19 (26.0%)	26 (34.2%)	45 (30.2%)
Dorset	29 (39.7%)	30 (39.5%)	59 (39.6%)
BCP	25 (34.2%)	20 (26.3%)	45 (30.2%)
Age, mean (SD)	42.21 (5.46)	43.09 (6.71)	42.66 (6.13)
Number of children per family [a], mean (SD)	1.42 (0.54)	1.51 (0.73)	1.47 (0.64)

[a] n = 45 for MBT group, n = 45 for TAU group, n = 90 for whole sample.

Note: TAU = treatment as usual; MBT-PP = mentalization-based therapy for parenting under pressure; BCP = Bournemouth, Christchurch and Poole; SD = standard deviation

Demographic data for CYP at baseline by condition is presented in Table 8. In most cases (88.6%), the IPC was occurring between CYPs' biological parents.

Table 8: Children and young people baseline sociodemographic characteristics (n = 132)

Characteristic	MBT-PP group (n = 64)	TAU group (n = 68)	All CYP
Sex, n (%)			
Male	30 (46.9%)	25 (36.8%)	55 (41.7%)
Female	34 (53.1%)	43 (63.2%)	77 (58.3%)
Ethnicity, n (%)			
White British	53 (82.8%)	57 (83.8%)	110 (83.3%)
White Irish	0 (0.0%)	0 (0.0%)	0 (0.0%)
Any other White background	6 (9.4%)	8 (11.8%)	14 (10.6%)
Mixed/multiple ethnic groups	5 (7.8%)	1 (1.5%)	6 (4.5%)
Asian or Asian British	0 (0.0%)	1 (1.5%)	1 (0.8%)
Black or Black British, Caribbean or African	0 (0.0%)	0 (0.0%)	0 (0.0%)
Other ethnic group	0 (0.0%)	1 (1.5%)	1 (0.8%)
Special Educational Needs / Disabilities, n (%)			
Yes	18 (28.1%)	21 (30.9%)	39 (29.5%)
No	46 (71.9%)	47 (69.1%)	93 (70.5%)
Looked-after child/previously looked-after child, n (%)			
Yes	2 (3.1%)	2 (2.9%)	4 (3.0%)
No	62 (96.9%)	66 (97.1%)	128 (97.0%)
Eligible for free school meals, n (%)			
Yes	13 (20.3%)	26 (38.2%)	39 (29.5%)
No	80 (78.1%)	42 (61.8%)	92 (69.7%)
Not disclosed	1 (1.6%)	0 (0.0%)	1 (0.8%)
Refugee or asylum seeker, n (%)			
No	64 (100.0%)	68 (100.0%)	132 (100.0%)
English as a second language, n (%)			
Yes	7 (10.9%)	6 (8.8%)	13 (9.8%)
No	57 (89.1%)	61 (89.7%)	118 (89.4%)
Not disclosed	0 (0.0%)	1 (1.5%)	1 (0.8%)
Interparental conflict between:, n (%)			
Biological parents	59 (92.2%)	58 (85.3%)	117 (88.6%)
Biological mother and stepfather	3 (4.7%)	8 (11.8%)	11 (8.3%)
Biological mother and other	2 (3.1%)	0 (0.0%)	2 (1.5%)
Adoptive parents	0 (0.0%)	2 (2.9%)	2 (1.5%)
Age, mean (SD)	10.56 (2.09)	10.99 (1.87)	10.78 (1.99)

Note: TAU = treatment as usual; MBT-PP = mentalization-based therapy for parenting under pressure; SD = standard deviation

While not directly comparable due to the age range of our sample, it is noteworthy that the sample represents a slightly higher proportion of CYP eligible for free school meals (29.5%) than the national average for all pupils (23.8%; DfE, 2024b). In our sample, 3% of children were reported as being (or having previously been) looked after, whereas the national proportion of children who were looked after by a local authority in 2024 was 0.7%, while 0.26% ceased to be looked after during the previous year (DfE, 2024a). Our sample also had a much higher proportion of children with SEND (29.5%) than the national average of pupils with an education, health and care (EHC) plan (4.3%; DfE, 2023) or without an EHC plan but receiving SEND support (13.6%). It should be noted, however, that these are not directly comparable, as parents reported on their children's SEND by responding yes/no to 'Does your child have any special educational needs or disabilities?' rather than the presence or absence of an EHC plan or SEND support in the absence of an EHC plan.

Evaluation feasibility

To appraise the feasibility of the trial design and processes, we have examined issues pertaining to referral pathways and recruitment, acceptability, attrition, power calculation estimate updates, data collection methods, data linking processes, support for the theory of change and the relationship between parents' SDQ reports.

Recruitment

As discussed in the introduction (and illustrated in Table 5), a significant task in the set-up phase for this project was the creation of referral pathways through which families could be recruited to take part. This process was not straightforward and required an agile problem-solving approach from the LA staff.

While formal data were not collected on referral sources, GLs reported that referrals to the project came from a range of sources. Notably, these differed somewhat according to each LA. BCP referrals largely came from targeted family support, early help, multi-agency safeguarding hub first response and schools, with a very small number coming from Dorset, Cafcass (which is the name for the Children and Family Court Advisory and Support Service), and services for young carers. For Bristol, referrals mainly came from schools, self-referrals, the edge of care team ('family help') and promotion events, with a handful also coming from autism support groups, social prescribing and GPs. For Dorset, the vast majority of referrals came from schools.

To capture the learning from this extensive work package, we included questions in interviews with the GLs about the challenges they faced in creating and embedding the referral pathways. The lessons learned serve to inform recruitment and referral processes for the efficacy phase, but they also speak to the challenge of supporting families with an issue that falls across and between the various spheres of service provision and governance that interact with family life.

RQ 1: To what extent are the referral pathways working?

To address this RQ, we analysed qualitative data from the GLs about their experiences of creating and embedding the referral pathways. We also examined the number of referrals to the project, the proportion of referrals that were eligible and the way in which these figures have changed over the course of recruitment so far. Initially, the GLs experienced some challenges in setting up referral pathways.

Two GLs noted that in the beginning, there was a lack of co-designing for promotional materials and that the way the materials were worded caused challenges in recruiting participants and organisations to participate in the trial or the development of referral pathways:

> *'The key challenges there, there was no co-production'.*

> *'The wording of the promotional materials and the way the project was started, we were trying to reach quite a very niche section of parental conflict'.*

Another issue, according to two of the GLs, was that due to the rapid set-up and the originally planned start-up period being reduced, the process of trying to develop relationships with external organisations to create referral pathways was challenging and time-consuming:

> '*I think because the start-up time was reduced so much, we weren't able to put a lot of things in place initially'.*

> *'Spending a lot of time promoting it and actually [making] that building of that professional relationship and understanding is probably the biggest challenge to get in'.*

To overcome those barriers, external organisations needed to be contacted repeatedly and reminded about the project many times, which led to building relationships with professionals:

> *'So it's been that sort of constant revisiting the same groups of people, the same areas, the same professionals, to sort of keep it fresh in people's minds, so that takes a lot of time, but it's proving to be worthwhile because every time we do that, every time we run a training session in an area or we go and visit a certain area in [redacted], we find within a week or two we are getting referrals from them'*.

Flagging the potential reduction in the workload of social workers in the area also seemed to help:

> *'Go and see them all and talk about SIPCo [Support for InterParental Conflict] and talk about how to refer into it, what it's all about, what my role is in supporting them and sort of almost saying that we're here to sort of reduce your workload, which family workers and social workers have really appreciated'.*

The process of recruiting parents subsequent to referral was more complex and took longer than the GLs had originally anticipated. The recruitment process was also slowed down by the time taken to complete the initial paperwork and assessments. This delay can be frustrating, especially when GLs find themselves waiting for a final piece of paperwork or a school appointment for a child, which can cause them to miss their targets for the month:

> *'We've got families that are perfect, and they want to be involved, but it just takes a really long time to get all of that initial sort of paperwork, measures, get all of that done before they can actually start'.*

Sometimes, lengthy recruitment periods would end with a family being ineligible rather than recruited. Reasons for this that were identified by all GLs were IPC-related barriers between parents, court proceedings and other safeguarding concerns.

However, the GLs developed strategies that worked, including methods for contacting parents and providing referrers with all the information they needed. As a result, they saw an improvement in recruitment. Two GLs noted that adapting to participant needs, such as meeting them in person or over the phone, was helpful:

> *'If they didn't pick up an email, [they] will receive a phone call from me, and that seems to be getting the ball rolling a lot quicker'.*

> *'I like to meet with [families] in person on the first meeting just to help build up that relationship and be able to have an open conversation'.*

All GLs noted that making initial contact with participants quickly after they had been referred was vital to the success of referral pathways:

> '*I think at LA [redacted], we try and have a seven working day turnaround, so from first contact, be that a referral or self-referral, to actually starting the programme, whichever it might be, so we aim for seven working days. It's not always possible, but that's kind of what we go for, so we either get that or close to it. I think families have really appreciated that because a lot of the support that's out there has huge waiting lists'.*

After families had been screened for eligibility and provided consent to participate, it took a median of eight days (interquartile range [IQR] = 18 days) to collect baseline survey data from all participating family members. This differed between LAs, with Dorset having a median of six days (IQR = 12 days), BCP having a median of 13 days (IQR = 16 days) and Bristol having a median of 17 days (IQR = 38 days).

The GLs were confident that future recruitment would be steady, reporting expectations of an average of six referrals each per month:

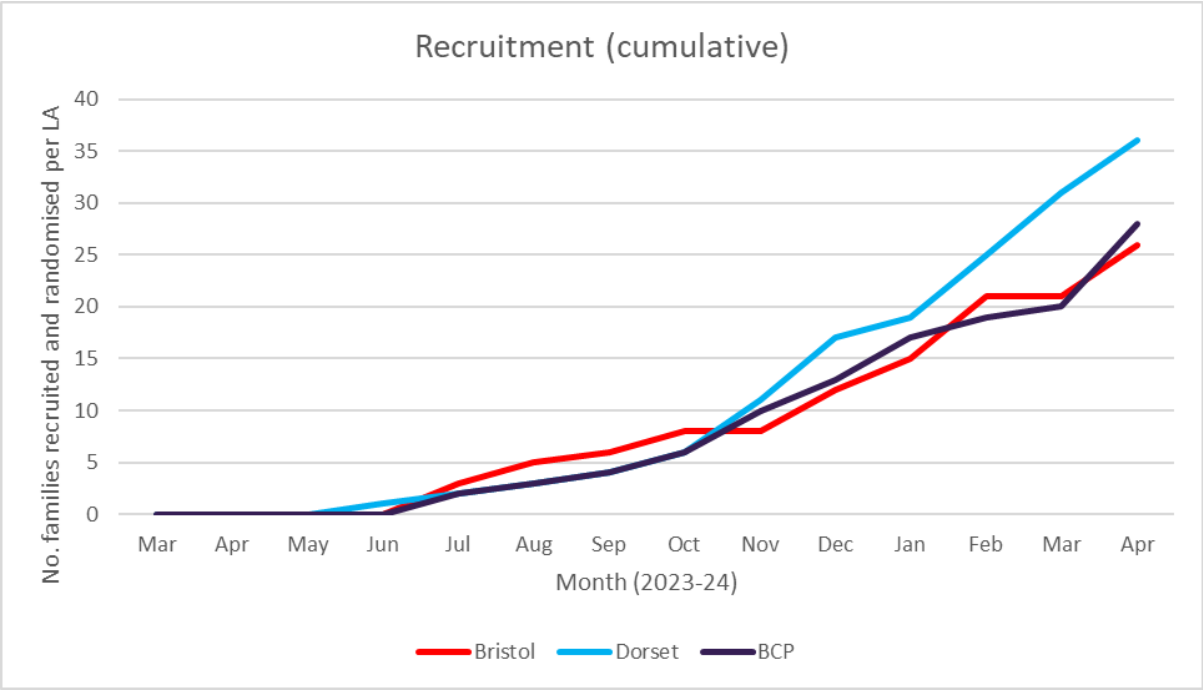
> *'We tend to have, I think, since October, we have had sort of between five and seven referrals every month that have gone through to allocation'.*

> *'I would say six is probably a steady number'.*

While the referral pathways took a long time to create and embed, the recruitment numbers have seen a steady increase since the first referral into the project (Figure 2). There have been differences in recruitment rates between the sites; hence, we have made efforts to share learnings through dialogue in regular meetings and presentations from the highest recruiting site. Weekly meetings between all three GLs, TR and evaluators have ensured that the GLs have as much support as possible in working out what works to generate referral pathways in their localities, sharing best practices and tackling queries about particular families' eligibility or complex circumstances. All three GLs reported that they found these meetings useful:

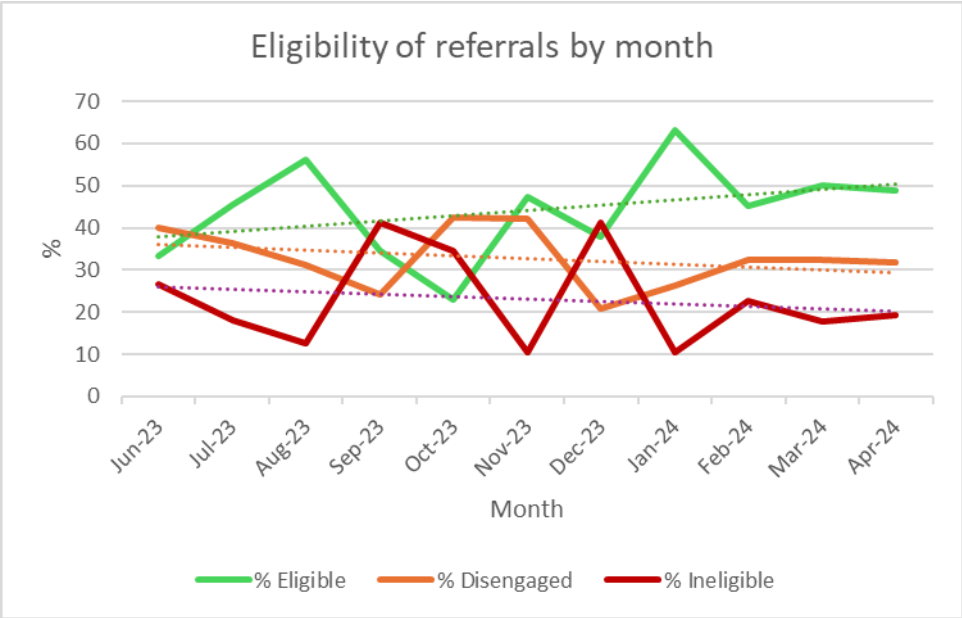
> *'And we've had our weekly meetings with the gateway leads [SHU trial manager, redacted] and [Tavistock staff, redacted], which has been really helpful, so we can sort of discuss any families we've got, any queries we've got about things, rather than just sort of emailing them or talking to them one on one, being able to do it as a group discussion has been really useful; so much more has come out of it'.*

Figure 2: Cumulative recruitment since the project opened to referrals



Recruitment numbers are indicative of those referred to the project who consented to participate, were deemed eligible, completed baseline measures and proceeded to be randomised and allocated to one of the trial arms. But to fully appraise the participant pipeline, it is important to also examine the referrals to the project and the proportion of those who are then eligible. These data show some evidence that referrers have become more consistent in terms of the quality of their referrals in the past three months, with the proportion of referrals deemed eligible vs ineligible vs disengaged (i.e. stopped responding to contact attempts or explicitly changed their minds about participating) becoming more stable. Figure 3 shows that from February 2024 onwards, the proportion of eligible referrals was around 50%, the proportion of ineligible referrals was around 20% and the proportion of referrals that later disengaged was around 30%. Trendlines also show the proportion of eligible referrals increasing since the start of the project, while the proportions of ineligible and disengaged referrals decreased.

Figure 3: Proportion of referrals that meet eligibility requirements



Acceptability and attrition

An important aspect that defines whether a full-scale efficacy RCT is feasible is whether the research processes and interventions are deemed acceptable by participants and whether there are any points at which problematic attrition might be indicative of processes which are unacceptable to participants.

RQ2: To what extent are the referral pathways, consent and randomisation procedures acceptable to participants (indicated by dropout rates at these points)?

To answer RQ 2, we can examine Figure 1, the participant flow diagram. We also gained important insights from our qualitative interviews with the GLs on the acceptability of the project to parents.

Of the 267 referrals received into the project before 1 May 2024, 77 (28.8%) declined to participate, and 190 (71.2%) expressed an interest and progressed to screening. Of those, 69 (36.32% of those expressing an interest; 25.8% of referrals received) were then deemed ineligible. This indicates that, broadly, the project is acceptable to parents, as >70% of those invited to be screened for eligibility go on to engage with the screening process. However, it is possible that this conversion rate of referral to screening for eligibility could be improved. We examined the qualitative data collected from GLs to understand the factors underlying this conversion rate and how improvements could be made.

As mentioned above, GLs talked about the importance of the speed with which they react to first contact with parents. Additionally, GL feedback instigated a rebranding of the project to avoid the use of the word 'conflict', which was identified as a barrier for many parents:

> *'When you think about the word conflict, they think of domestic abuse; it's just the way that particularly social workers' brains are kind of wired. That conflict is that high-level abuse. And the same for parents … parents don't want to say they're in conflict because they are thinking that if they admit that, they're going to get reported to social services or that they are doing something wrong … if it's branded and encouraged that, "Would you like to improve your communication with your partner or ex-partner?" A lot more families and parents are going to go, "Yes, actually, I'll be up for that"'.*

RQ 3: Are there any signs of problematic attrition (e.g. that might indicate that the research processes or interventions are not acceptable to participants)?

The participant flow diagram (Figure 1) also allows us to answer RQ3. So far, the only point after randomisation that we see participant dropout is at post-intervention data collection. This indicates the acceptability of the research processes and interventions up to that point. At post-intervention data collection, however, attrition affects 20% of the MBT arm (four families) and 23.5% of the TAU arm (four families). Importantly, however, it is possible that some of these data may yet be collected, as the project is currently 'live', meaning that these numbers are continually being updated. Also, it should be acknowledged that the overall numbers here are small because most recruited families are still undergoing their interventions. This is not to downplay the significance of attrition at this point in the research process, as it is this data collection point at which the primary outcome variable is gathered. We are also aware that both of the red-rated progression criteria are related to this attrition point. We have therefore considered how to restructure the incentives to make completion of this data point more rewarding.

As per the research protocol, voucher incentives were used to thank participants (parents and CYP) in both arms of the trial for questionnaire completion. To minimise attrition from the research processes (completion of measures), we used a structured incentive system. Vouchers for questionnaire completion were provided after each data collection point, and a further voucher was provided as a bonus for completion of all three time points after the final questionnaire. We propose that moving forward to efficacy, we will offer the bonus voucher at the completion of the post-intervention questionnaire rather than at the three-month follow-up. This will make the post-intervention questionnaire twice as rewarding as baseline and follow-up. While it is possible that this will reduce the motivation to complete the follow-up questionnaire, the point is moot if we are losing participants earlier in the pipeline.

We also propose that we utilise a behavioural-science-informed approach to the wording of emails that are sent to ask participants to complete questionnaires. We will seek to optimise the wording of these in line with the EAST framework (Behavioural Insights Team, 2014), which posits that communications should be easy, attractive, social and timely to enable maximal compliance. To make it easy for participants to know what they should do at each time point, a series of email templates will be provided to the LAs. The emails will clearly provide the relevant survey link(s) near the start of the email so that they are not missed by the participants. The emails will use attractive formatting and emphasise the shopping voucher incentives that the participants will receive as thanks for their time completing the surveys. The GLs will personalise the start and end of each email based on their ongoing social relationships with the participants. The emails will also ask the participants to complete the survey(s) in a timely manner – within one week.

All GLs reported that the trial processes were broadly acceptable to participants. This included participant information sheets, consent forms and questionnaires:

> *'Parents are quite happy to do the consent form … they're quite happy to consent'.*

> *'In terms of doing the measures, the questionnaires as it were, the children quite enjoy it, I think; they like the fact [that] they get out of class for a few minutes, so that's always nice'.*

All GLs also reported that some families were a bit disappointed when they were randomised to their least preferred condition, and this required some additional explanation and support:

'And so, then you have to re-go through everything, which can be quite tricky because parents might have in their head what they would like to get out of the randomisation and what their expectation is, and it's not always that'.

On balance, however, the GLs felt that the research processes and intervention were acceptable to the participants, and the positive feedback received from those who have completed TAU has been incorporated into new advertising for the project.

RQ 4: How should the estimates used for the sample size calculation be adjusted in light of the data?

This trial is powered on the number of CYP (aged 8–14) within eligible families, as the maternal total difficulties SDQ score (for each CYP aged 8–14) is our primary outcome measure.

The minimum detectable effect size (MDES) was calculated using Equation 1 (below), adapted from Spybrook, Kelcey and Dong (2016). We estimated the MDES for our design using this equation and checked this using the PowerUp! software (Dong et al., 2015).

Equation 1

$$MDES_{2LCRT} \sim M_{J-m-2}\sqrt{\frac{1}{P(1-P)}}\sqrt{\frac{ICC(1-R^2_{Fam})}{J}+\frac{(1-ICC)(1-R^2_{CYP})}{nJ}}$$

A number of parameters affect the MDES in a clustered randomised trial, and those in bold below are the estimates that will be adapted using the pilot data:

- n = number of children per family (**estimated as 2**)
- J = number of families
- P = proportion of families allocated to the intervention group (**0.50**)
- m = number of (level 2) covariates used (which will include group membership, family-level pre-test and all variables used for minimisation, ~**11** variables)
- M_{J-m-2} is the group effect multiplier value of the t-distribution for a two-tailed test with alpha = 0.05 and beta = 0.80
- ICC is the family-level ICC (intra-class correlation); which is the proportion of of the outcome at level 2 (between family variance). This is unknown but **estimated at between 0.01 and 0.15 (to be updated with pilot data)**
- R^2_{CYP} = proportion of within-family child-level variance that is reduced by covariate(s) – child-level explanatory power and R^2_{Fam} = proportion of between-family variance that is reduced by covariate(s) – family-level explanatory power. **These will be updated using pilot data**. For these a priori MDES estimates, we have assumed that $R^2_{CYP} = R^2_{Fam}$ and allowed the values to vary between 0.25 and 0.49 (based on an **assumed pre-post-test correlation of between 0.50 and 0.70, to be updated with estimates from pilot data**).

This power calculation indicated that the recommended sample for the trial (pilot + efficacy) was 350 families (with 700 CYP) to be able to detect an effect size of 0.20 or higher as statistically significant ($p < 0.05$, two-tailed) with a statistical power of 0.80 or higher.

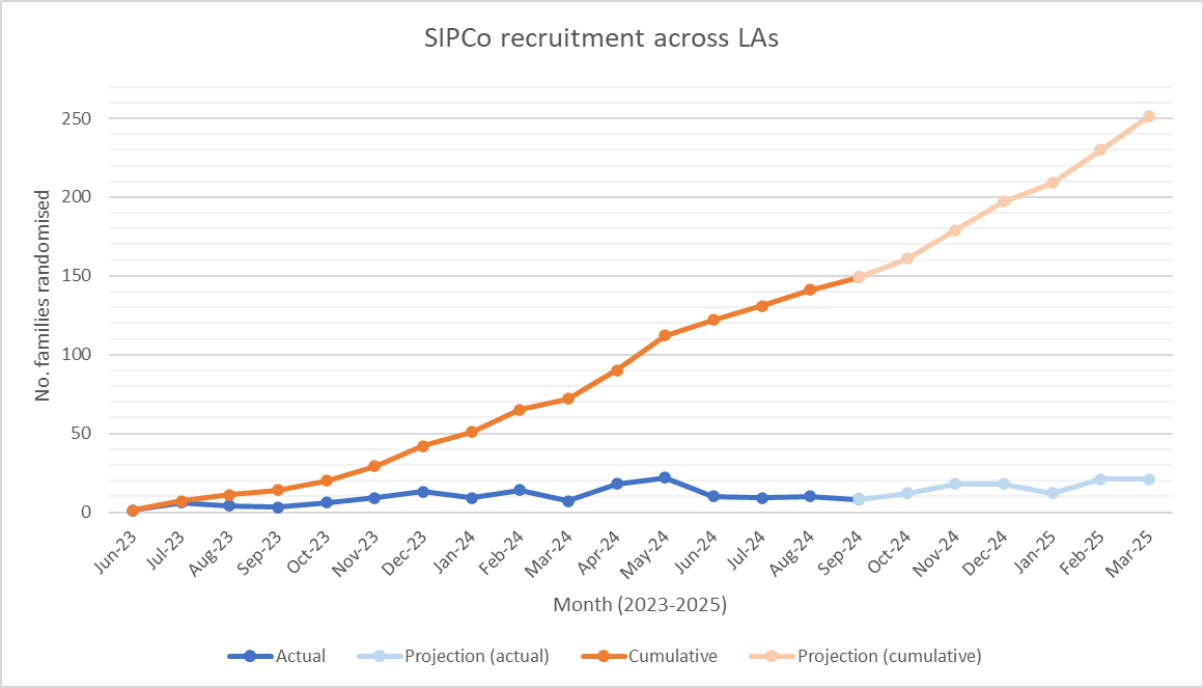
To inform the development of the Statistical Analysis Plan (SAP), we used the data from the pilot phase to update the estimates on which the power calculation was based. The changes made were:

- Cluster size, estimated as 2, was changed to 1.4
- ICC, estimated as 0.01–0.15, was changed to 0.27
- CYP explanatory power, estimated as 0.25–0.49, was changed to 0.75
- Family explanatory power, estimated as 0.25–0.49, was changed to 0.61

Within these changes, the factors that decrease sensitivity (i.e. raise the MDES) are having a notably higher ICC and a lower mean number of CYP per family. Conversely, the factors that increase sensitivity (i.e. lower the MDES) are the explanatory power being high at both the family and CYP levels.

The SAP sets out the full calculation, the results of which indicate that the recommended sample for this trial (pilot + efficacy) is 250 families (350 CYP), making the MDES estimate 0.17. This is on the basis that allowing for an attrition rate of up to 20% still provides an MDES of 0.19 SDs., assuming that the attrition is random. Based on recruitment data so far, the expectations of the GLs regarding how they think recruitment is likely to continue, and accounting for anticipated seasonal dips and peaks, we project an average of 17 referrals per month moving forward (Figure 4). We, therefore, propose that an extension to recruitment of three months (from the end of December 2024 to the end of March 2025) is required to ensure we meet the sample indicated by the power calculation.

Figure 4: Recruitment projection to end of project



RQ 5: Do the data collection methods work?

Our analyses of the data quality indicate that there are very few missing data. The percentage of missing data within the 25 SDQ items completed by parents at baseline was 0.04% (two missing responses out of

5,475 SDQ responses). At post-treatment, there were 0% missing responses (0 missing responses out of 1,150 responses).

We are also broadly satisfied with the reliabilities of the scales (presented in the data collection section above), most of which are good.

We also analysed the response times for the completion of baseline measures. Excluding surveys that took longer than one day to complete (i.e. they were not submitted properly by the participants after completion but instead auto-submitted by Qualtrics after a timeout period of one week), the average response time for parents to complete the baseline questionnaire was 33 minutes, and the average response time for CYP to complete the baseline questionnaire was nine minutes. We conclude from this that the data collection methods can be deemed reasonable in terms of time commitment from the participants.

Additionally, GLs reported that participants were generally comfortable with completing the measures:

> 'The baseline survey, parents are happy to do'.
>
> 'I haven't had any families that have said no'.

We therefore conclude that the data collection methods are working well.

Treatment fidelity

The pilot phase provided an opportunity to develop a prototype fidelity checklist and to explore the extent to which it could be used reliably to assess the fidelity of delivery of the MBT-PP intervention. The fidelity checklist was developed from the key components of MBT-PP, as described in the handbook (TR, n.d.), which were potentially observable during therapeutic interactions. These included:

1. Noticing and naming (observing patterns of interaction; checking out what you have noticed; agreeing on a name for an unhelpful pattern of interaction)
2. Mentalising the moment (being curious yourself; provoking curiosity in the parents; pausing, reviewing and marking the moment)
3. Generalising and considering change (widening the lens; considering constructive alternatives; planning action or considering specific small steps either parent could take)
4. Checking feeling states

In each case, raters were asked to state whether the component was 1) clearly observed and evidenced, 2) possibly observed but with some doubt or 3) definitely not observed. Raters recorded the time on the worksheet to aid moderation. The second part of the checklist listed ten skills that the therapist should demonstrate during an MBT-PP session on a 5-point Likert scale from 1 = 'Not at all demonstrated' to 5 = 'Very clearly demonstrated'.

Three raters from the research team (one with a background as a therapist, one a registered health psychologist and one who has published on therapeutic processes) independently listened to a sample audio recording of an MBT-PP session delivered during the pilot phase and used the prototype fidelity checklist to rate the session. They met to discuss the rating and used this process to make minor changes to the checklist to ensure consistency in what was and was not being included under each component, whether there was agreement about what had been observed and whether Likert scale ratings were broadly similar. The same three raters used a revised version of the checklist to rate the second and third audio recordings for fidelity, and further changes were made to the checklist. The final version of the fidelity checklist is provided in

Appendix F. It will be used during the main trial phase to explore the extent to which therapists are proficient and consistent in their delivery of MBT-PP. Mean scores will be calculated from each sample audio recording per therapist to appraise fidelity and contextualise the efficacy findings.

RQ6: Do data linking processes work?

Once family details were provided to the SHU research team by the GLs, the research team systematically allocated each family a unique family ID, and each family member was allocated a unique participant ID. These IDs contained information about which LA the family was from, the referral number from that LA and, for the participant IDs, which family member each individual was. Baseline surveys automatically generated a random 10-digit number, which the participants were provided with and asked to enter into the two subsequent surveys to assist with linking data from each family member across the three time points (baseline, post-treatment and follow-up). These random 10-digit numbers were used as a failsafe to mitigate situations in which more than one participant had the same name or where participants entered their names slightly differently at different time points (e.g. used shortened forms of their first name in one survey but not in others or included typographical errors). Survey responses were tagged with participant IDs, using automated processes (data merges in Excel) where possible and manually in cases where the random ID had not been entered by the participant (this happened in nine of 39 [23.1%] parent post-treatment surveys and 0 of six [0%] parent follow-up surveys completed by 1 May 2024). Therefore, all data linking processes worked as intended, with data from each participant being identifiable to the research team and linked to that of the other relevant family members and across time points.

Evidence of promise

To assess evidence of promise, we examined the descriptive statistics for all variables by trial arm. We also asked in the interviews with the GLs whether they had received any informal feedback about the interventions.

RQ7: Is there any early evidence supporting the theory of change?

To examine whether there is any evidence of promise emerging from the pilot data, we inspected the descriptive statistics for all outcome variables (Tables 9–12), although caution should be used in drawing any conclusions, given the small sample size involved and the lack of follow up data.

Table 9: Child-specific measures from mothers (n = 45)

Measure	MBT-PP (n = 23)				TAU (n = 22)			
	Baseline		Post-intervention		Baseline		Post-intervention	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
SDQ total difficulties	16.17	6.98	15.09	6.74	15.64	7.87	14.77	7.08
SDQ emotional symptoms	5.08	2.98	4.52	3.01	5.23	2.60	4.41	2.81
SDQ conduct problems	2.65	1.80	2.22	1.88	2.64	2.36	2.41	2.20

SDQ hyperactivity/inattention	5.78	3.00	5.39	2.73	4.82	2.92	4.77	2.72
SDQ peer problems	2.65	2.48	2.96	2.84	2.95	2.50	3.18	2.44
OPS overt hostility [a]	20.52	7.17	25.61	6.50	22.91	7.27	26.73	7.47
PRFQ pre-mentalising modes	1.91	0.68	2.12	1.19	1.79	0.73	1.95	1.05
PRFQ certainty about mental states	4.01	1.37	4.22	1.12	4.19	1.43	4.19	0.89
PRFQ interest and curiosity about mental states	6.42	0.56	6.01	1.29	5.92	0.61	5.84	1.02
PS-8[b]	3.10	1.07	2.97	1.09	3.05	1.11	2.70	1.06

[a] Higher scores on this measure indicate lower overt hostility. [b] Higher scores on this measure indicate more dysfunctional parenting strategies.

Note: MBT-PP = mentalization-based therapy for parenting under pressure; TAU = treatment as usual; SDQ = Strengths and Difficulties Questionnaire; SD = standard deviation; OPS = O'Leary-Porter Scale; PRFQ = Parental Reflective Function Questionnaire; PS-8 = Parenting Scale Short Form

Table 10: Child-specific measures from fathers (n = 17)

Measure	MBT-PP (n = 7)				TAU (n = 10)			
	Baseline		Post-intervention		Baseline		Post-intervention	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
SDQ total difficulties	11.71	9.25	14.43	9.40	14.70	6.53	15.30	6.11
SDQ emotional symptoms	3.14	3.80	5.00	3.92	4.40	2.76	3.80	2.53
SDQ conduct problems	2.57	2.30	1.86	1.95	2.70	1.49	2.90	1.79
SDQ hyperactivity/inattention	4.29	3.20	4.57	3.64	4.90	2.38	5.30	2.00
SDQ peer problems	1.71	1.89	3.00	2.38	2.70	2.75	3.30	3.23
OPS overt hostility [a]	24.57	6.19	26.84	5.43	22.80	6.68	27.40	5.74
PRFQ pre-mentalising modes	1.79	0.80	2.48	0.80	2.15	0.55	1.72	0.35
PRFQ certainty about mental states	3.50	0.82	3.93	0.69	3.25	0.58	3.50	0.62
PRFQ interest and curiosity about mental states	5.29	1.26	4.65	1.61	5.30	1.06	5.53	0.88
PS-8[b]	3.06	0.49	2.71	0.50	3.87	1.26	3.37	0.80

[a] Higher scores on this measure indicate lower overt hostility. [b] Higher scores on this measure indicate more dysfunctional parenting strategies.

Note: MBT-PP = mentalization-based therapy for parenting under pressure; TAU = treatment as usual; SDQ = Strengths and Difficulties Questionnaire; SD = standard deviation; OPS = O'Leary-Porter Scale; PRFQ = Parental Reflective Function Questionnaire; PS-8 = Parenting Scale Short Form

Table 11: Parent-specific measures (n = 44)

Measure	MBT-PP (n = 22)				TAU (n = 22)			
	Baseline		Post-intervention		Baseline		Post-intervention	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Dimensions of Anger Reactions – Revised	8.32	4.44	4.95	3.31	7.73	5.86	6.68	4.49
Emotional Adaptation Relationship Dissolution Assessment [a]	1.88	1.09	2.37	1.14	3.05	0.58	2.59	1.13

[a] n = 6 in MBT-PP condition at baseline, n = 11 in TAU condition.

Note: MBT-PP = mentalization-based therapy for parenting under pressure; TAU = treatment as usual

Table 12: Children and young people measures (n = 32)

Measure	MBT-PP (n = 21)				TAU (n = 11)			
	Baseline		Post-intervention		Baseline		Post-intervention	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Stirling Children's Well-being Scale	40.00	8.83	40.14	6.77	38.09	7.87	41.09	7.35
Perceptions of Interparental Conflict Intensity/Frequency Scale	41.76	7.93	39.33	7.00	40.73	5.10	39.27	5.26

Note: MBT-PP = mentalization-based therapy for parenting under pressure; TAU = treatment as usual

Elements of the logic model were supported or not, by these data in the following ways:

Compared to baseline, there were changes post-intervention in measured outcomes, including:

- Capacity to mentalise: for both mothers and fathers in the MBT-PP condition, scores on two mentalising subscales increased (pre-mentalising modes and certainty about mental states), and scores on one subscale decreased (interest and curiosity about mental states). For mothers in the TAU condition, scores on one mentalising subscale increased (pre-mentalising modes), scores on one subscale did not change (certainty about mental states) and scores on one subscale decreased (interest and curiosity about mental states). For fathers in the TAU condition, scores on two mentalising subscales increased (interest and curiosity about mental states and certainty about mental states), and scores on one subscale decreased (pre-mentalising modes).
- Anger: for both the MBT-PP and TAU, there was a reduction in anger.

- Higher adaptation to separation (where relevant): this increase was evident. The opposite pattern was apparent in TAU.
- Lower parent-reported conflict: this reduction was evident for mothers and fathers, and a similar pattern occurred in TAU.
- Lower child-reported perception of conflict: this reduction was evident, and a similar pattern occurred in TAU.
- Higher child well-being: there was no discernible change in MBT, but an increase was evident in TAU.
- Lower SDQ difficulties: this reduction was evident for mothers, but the opposite pattern occurred for fathers. Similar patterns occurred in TAU.

While there are no defined feedback opportunities for participants to let the GLs know how they got on with MBT, one GL reported receiving positive responses:

> *'He feels like he has learned a lot of skills to be able to communicate better … he has found [it] really helpful, and he is going to be able to apply it to work as well, so work relationships and not just in his personal life, so he's been very grateful for that'.*
>
> '*We've had parents that have done MBT and treatment as usual [who] have said it's actually really helped their relationship with their child as well'.*

Overall, it is unclear whether there is early evidence of promise in support of the theory of change, in which MBT is expected to increase parent mentalising, which, in turn, decreases their anger expression and increases their adaptation to their separation where relevant. IPC is consequently expected to reduce and so is the CYP-perceived IPC, which, in turn, should reduce the CYP total difficulties score. However, at this stage, the small sample size involved means that it is very unlikely that any conclusions regarding support for the theory of change can be made at this point.

RQ8: What is the relationship between parents' SDQ reports?

To assess the relationship between mothers' and fathers' SDQ scores (in order to inform the strategy for handling missing data on the primary outcome variable of mothers' SDQ scores by appraising the value of fathers' SDQ scores as a proxy), we examined correlations at the level of CYP (Table 13).

Table 13: Correlations between mothers' and fathers' SDQ scores

Measure	Correlation between mothers' and fathers' SDQ scores					
	MBT-PP group		TAU group		Whole sample	
	Baseline (n = 41)	Post-intervention (n = 7)	Baseline (n = 46)	Post-intervention (n = 10)	Baseline (n = 87)	Post-intervention (n = 17)
SDQ total difficulties	0.49**	0.63	0.70**	0.84**	0.62**	0.72**
SDQ emotional symptoms	0.41**	0.16	0.63**	0.75*	0.53**	0.38
SDQ conduct problems	0.44**	0.71	0.59**	0.57	0.51**	0.63**
SDQ hyperactivity/inattention	0.39**	0.97**	0.53**	0.31	0.45**	0.66**
SDQ peer problems	0.62**	0.40	0.70**	0.87**	0.66**	0.74**

* Correlation is significant at the 0.05 level (two-tailed). ** Correlation is significant at the 0.01 level (two-tailed).

Note: MBT-PP = mentalization-based therapy for parenting under pressure; TAU = treatment as usual; SDQ = Strengths and Difficulties Questionnaire

It is notable that the strength of the correlation increases from baseline to post-intervention across all but one subscale (emotional symptoms), where it reduces. This may be indicative of parents reaching a greater agreement regarding their children's difficulties after receiving the interventions, but it is interesting to consider why the emotional symptoms subscale shows the opposite pattern. Given the very small sample size at post-intervention, these findings may not be reliable and may change with increasing statistical power. Nonetheless, the literature suggests that parent reports can differ (Bergström and Baviskar, 2021), and the delivery partners expected that differences between mothers and fathers would be higher than average in a sample with high IPC because disagreement about children's well-being is often a topic of IPC.

Note that for this reason, it is not appropriate to substitute mothers' SDQ scores with fathers' SDQ scores in the primary analysis in cases in which we have missing data for mothers. In non-traditional families (two mothers, two fathers or non-binary parents), our approach is to randomly select which SDQ to treat as primary.

Cost information

The pilot phase has revealed some limitations in the cost data collection processes; therefore, we regard this as a work in progress to be rectified for the efficacy phase. Based on the data we currently have (which does not include staff training costs), we estimate the set-up and recurring costs as follows:

Set-up costs

Set-up costs include training therapists to deliver MBT-PP and marketing/communications costs associated with advertising the project.

Training costs total to 30 April 2024 = £3,000 for staffing/running the training course and £1,500 per therapist x 5 therapists for attending. Total training spend = £13,500

Marketing/communications costs total to 30 April 2024 = £84,717 for 90 families. Forty-five families were randomised to MBT-PP, meaning £91 was spent on advertising per family recruited.

Recurring costs

Recurring costs included therapist time to deliver MBT-PP, supervisor time to supervise MBT-PP therapists, overheads (office running costs), support staff, ICT, phones and equipment.

Delivery of therapy sessions	£138,908
Supervision	£31,144
Overheads (office running costs)	£30,939
Other support staff	£24,822
ICT, phones and equipment	£9,940
Total spend on intervention delivery	£235,753

Total number of therapy sessions delivered:

Twenty families completed MBT-PP (10 sessions per family); 25 were still in treatment (estimate of five sessions per family).

= 325 sessions delivered.

Recurring cost total to 30 April 2024 = £235,753 for approximately 325 sessions of MBT-PP delivered. Cost per session = £725. Cost per 10-session treatment (per family) = £7,250. Based on the mean of 1.4 CYP per family, this equates to a cost of £517 per CYP per session, or £5178 per CYP per 10-session treatment.

We are working with the delivery team to achieve more complete and accurate data, both for the pilot phase and moving forward to efficacy.

Conclusion

Evaluator judgement of evaluation feasibility

It is our view as the evaluation team that there would be significantly more value in proceeding to full efficacy than halting the research following the pilot phase. Considerable challenges were faced and overcome in the set-up phase for this project, and progressing to efficacy would capitalise on the huge achievements of the GLs to set up referral pathways for receiving support for IPC in their LAs. Implementing such change is a vast undertaking and perhaps one that was underestimated by all parties at the beginning of this work. Now that these pathways have been created and are beginning to be embedded and fully utilised by referrers, the foundational work is in place for a successful efficacy phase. Current evidence suggests that the trial processes are acceptable to both participants and the GLs. Regarding delivery, MBT-PP is being delivered in a timely manner, and informal feedback to the GLs suggests that it is positively received by parents.

The pilot phase has indicated that the majority of the progression criteria have been met (rated green, see Tables 1–3), and the following mitigations have been proposed for the minority of criteria which are rated amber or red:

To increase recruitment:

1. GLs will be encouraged to highlight to professionals who may refer to the project (such as social workers and family support workers) that doing so could lighten their workload because the project could potentially offer support that they themselves do not have to provide.
2. The project will be re-branded to focus on improving parental communication rather than reducing parental conflict. This is expected to appeal more to parents and help professionals understand who the project is targeting (i.e. not cases of domestic abuse).
3. GLs/administrators will be encouraged to adopt target timescales for getting in touch with parents after initial contact, as this is thought to have been helpful in raising recruitment in one area.

Of note, while the number of families recruited at the end of April was in the amber category, with 90 families recruited, the current sample size as of 31 May is 110 families (159 CYP), which is in the green category.

To reduce attrition:

1. Incentives will be re-structured so that the post-intervention questionnaire (for parents and CYP) will be more heavily incentivised by bringing the bonus voucher forward.
2. LAs will be provided with behavioural-science-informed email templates to support their efforts to retain parents in the trial.

Interpretation

This internal pilot for a pragmatic cluster RCT has demonstrated that the main stage evaluation is feasible and should go ahead as planned.

Project implementation progression criteria were all green except one. In summary, the therapy sessions were recorded, most MBT-PP clients attended enough therapy sessions (although this was amber rather

than green), MBT-PP therapists received enough training and supervision, and they had sufficient capacity to work with clients in a timely manner. Sufficient referrals met the eligibility criteria. While referral pathways took time to generate and embed, they are now working well. Mitigations to guard against low recruitment in the efficacy phase include reminding the GLs what works best in terms of pitching to professionals and fast processing of new families and re-branding the project to parents in line with GL feedback to focus on improving communication rather than reducing conflict. Recruitment is a laborious process for the GLs, but they have well-practised processes in place. The system works best when the GLs are supported by dedicated administrators. Referral processes appear to be broadly acceptable to participants. While 29% of families referred to the project declined to participate (i.e. disengaged before/during the screening process), once deemed eligible, no families dropped out prior to randomisation and allocation, indicating that randomisation has good acceptability. Re-branding the project to parents may help to increase the number of referred families that agree to be screened for eligibility.

Evaluation measurement progression criteria were a mixture of amber, green and red. While almost enough families were recruited, the adjusted power calculation means that the project is on track to meet the recruitment needed for the efficacy study. Enough families that started MBT-PP completed the treatment protocol, and very nearly enough (only 0.8% away from green) primary outcome variable data (post-treatment maternal SDQ) were collected. Attrition may be problematic at post-intervention, so mitigations, which include increasing incentives for post-intervention measures and supporting GLs to use behavioural-science-informed approaches to prompt measure completion, have been suggested to improve this for the efficacy phase. While slight adjustments are recommended to increase recruitment and decrease attrition, these adjustments do not represent changes to data collection methods or interventions. This means that the aim of an internal pilot for the data to be combined with efficacy phase data is not undermined.

Measurement and findings progression criteria were all green. Randomisation processes worked without any problems. The pilot data revealed a possible imbalance in free school meal distribution, and the research team will monitor this moving forward. The baseline and post-treatment primary outcome measurement data are all of high quality. There is an effective mechanism in place to collect the outcome measures, in that the GLs are all in post and working to deliver data. Data have an extremely low proportion of missing responses. Scale reliabilities are generally good. Data linkage processes have also worked well without any problems.

Findings relating to outcome measures are mixed and somewhat unclear, possibly due to low statistical power. It is too early in the research to conclude any evidence in support of or against the theory of change for MBT-PP, although informal feedback suggests that parents find the interventions helpful.

Glossary

AB – Arguing Better

BCP – Bournemouth, Christchurch and Poole

CYP – Children and Young People

DAR-R – Dimensions of Anger Reactions – Revised

DAS-4 – Dyadic Adjustment Scale (four-item version)

EARDA – Emotional Adaptation to Relationship Dissolution Assessment

IPC – inter-parental conflict

GIRFC – Getting It Right For Children

GL – gateway lead

LA – local authority

MBT-PP – Mentalization Based Therapy for Parenting Under Pressure

MDES – minimum detectable effect size

OPO – OnePlusOne

OPS – O'Leary–Porter Scale

PIC-I/F – Perceptions of Interparental Conflict-Intensity / Frequency Scale

PMRs – pupil matching reference numbers

PRFQ – Parental Reflective Function Questionnaire

PS-8 – Parenting Scale Short Form

RCT – randomised controlled trial

SCWBS – Stirling Children's Well-being Scale

SDQ – Strengths and Difficulties Questionnaire

SHU – Sheffield Hallam University

TR – Tavistock Relationships

References

Acquah, D., Sellers, R., Stock, L. and Harold, G. (2017) *Inter-parental conflict and outcomes for children in the contexts of poverty and economic stress: implications aimed at improving outcomes for children* (Version 1). University of Sussex. Available at: https://hdl.handle.net/10779/uos.23476664.v1 (Accessed 11 September 2024).

Adamsons, K. and Pasley, K. (2006) Coparenting following divorce and relationship dissolution. In Fine, M. and Harvey, J. (eds.) *Handbook of Divorce and Relationship Dissolution*. Mahwah, NJ: Lawrence Erlbaum, pp. 241-262.

Altman, D. G. and Bland, M. J. (2005) Statistics notes: treatment allocation by minimisation, *BMJ*, 330, p. 843.

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

Assarroudi A., Heshmati Nabavi F., Armat M. R., Ebadi A. and Vaismoradi M. (2018) Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process, *Journal of Research in Nursing*, 23(1), pp. 42-55. Available at: https://doi.org/10.1177/1744987117741667.

Bateman, A. and Fonagy, P. (2010) Mentalization based treatment for borderline personality disorder, *World Psychiatry*, 9(1), pp. 11-15. Available at: https://doi.org/10.1002/j.2051-5545.2010.tb00255.x.

Behavioural Insights Team (2014) *EAST: four simple ways to apply behavioural insights*. London, UK: Behavioural Insights Team.

Bergström, M. and Baviskar, S. (2021) A systematic review of some reliability and validity issues regarding the Strengths and Difficulties Questionnaire focusing on its use in out-of-home care, *Journal of Evidence-Based Social Work*, 18(1), pp. 1-31. Available at: https://doi.org/10.1080/26408066.2020.1788477.

Brewin, W. and Garlick, M. (2023) *Reducing parental conflict programme 2018-2022: an evaluation of the effects of interventions on parental relationships and children*. Available at: https://assets.publishing.service.gov.uk/media/65cce78cc96cf300126a3779/reducing-parental-conflict-programme-2018_2022-evaluation-of-interventions.pdf (Accessed 31 May 2024).

Conger, R. D. et al. (1994) *Families in troubled times: adapting to change in rural America*. New York: Routledge.

Davies, P. T. and Cummings, E. M. (1994) Marital conflict and child adjustment: an emotional security hypothesis, *Psychological Bulletin*, 116(3), pp. 387–411. Available at: https://doi.org/10.1037/0033-2909.116.3.387.

Department for Education (2023) *Special educational needs in England, academic year 2022/23*. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/special-educational-needs-in-england (Accessed 31 May 2024).

Department for Education (2024a) *Children looked after in England including adoptions, reporting year 2024*. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/children-looked-after-in-england-including-adoptions (Accessed 23 January 2025).

Department for Education (2024b) *Schools, pupils and their characteristics, academic year 2022/23*. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics (Accessed 31 May 2024).

Department for Work and Pensions (2017) *Improving lives: helping workless families, analysis and research pack*. Available at: https://www.gov.uk/government/publications/improving-lives-helping-workless-families (Accessed 31 May 2024).

Department for Work and Pensions (2020, updated 2023) *Reducing parental conflict programme evaluation: report on early implementation*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1042489/reducing-parental-conflict-programme-evaluation-report-on-early-implementation.pdf (Accessed 31 May 2024).

Dong, N., Kelcey, B., Maynard, R. and Spybrook, J. (2015) *PowerUp! Tool for power analysis*. Causal Evaluation. Available at: https://www.causalevaluation.org/ (Accessed 31 May 2024).

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L. and Lancaster, G. A. (2016) CONSORT 2010 statement: extension to randomised pilot and feasibility trials, *British Medical Journal*, 355, i5239. Available at: https://doi.org/10.1136/bmj.i5239.

Goodman, R. (1997) The Strengths and Difficulties Questionnaire: a research note, *Child Psychology and Psychiatry and Allied Disciplines*, 38(5), pp. 581–586. Available at: http://dx.doi.org/10.1111/j.1469-7610.1997.tb01545.x.

Grych, J. H. and Fincham, F. D. (1990) Marital conflict and children's adjustment: a cognitive-contextual framework, *Psychological Bulletin*, 108(2), pp. 267-290. Available at: https://doi.org/10.1037/0033-2909.108.2.267.

Grych, J. H., Seid, M. and Fincham, F. D. (1992) Assessing marital conflict from the child's perspective: the children's perception of interparental conflict scale, *Child Development*, 63(3), pp. 558-572. Available at: https://doi.org/10.1111/j.1467-8624.1992.tb01646.x.

Hajek Gross, C. *et al.* (2024) Efficacy of mentalization-based therapy in treating self-harm: a systematic review and meta-analysis, *Suicide and Life-Threatening Behavior*, 54(2), pp. 317-337. Available at: https://doi.org/10.1111/sltb.13044.

Harold, G. T. et al. (2016) *What works to enhance inter-parental relationships and improve outcomes for children*. London, UK: Department for Work and Pensions.

Harold, G. T. and Conger, R. D. (1997) Marital conflict and adolescent distress: the role of adolescent awareness, *Child Development,* 68(2), pp. 333-350. Available at: https://doi.org/10.1111/j.1467-8624.1997.tb01943.x.

Harold, G. T. and Sellers, R. (2018) Annual research review: interparental conflict and youth psychopathology: an evidence review and practice focused update, *Journal of Child Psychology and Psychiatry*, 59(4), pp. 374-402. Available at: https://doi.org/10.1111/jcpp.12893.

Hertzmann, L., Abse, S., Target, M., Glausius, K., Nyberg, V. and Lassri, D. (2017) Mentalisation-based therapy for parental conflict–parenting together: an intervention for parents in entrenched post-separation disputes, *Psychoanalytic Psychotherapy*, 31(2), pp. 195-217. Available at: https://doi.org/10.1080/02668734.2017.1320685.

Hertzmann, L., Target, M., Hewison, D., Casey, P., Fearon, P. and Lassri, D. (2016) Mentalization-based therapy for parents in entrenched conflict: a random allocation feasibility study, *Psychotherapy*, 53(4), pp. 388-401. Available at: https://doi.org/10.1037/pst0000092.

Kliem, S., Lohmann, A., Mößle, T., Foran, H. M., Hahlweg, K., Zenger, M. and Brähler, E. (2019) Development and validation of a parenting scale short form (PS-8) in a representative population sample, *Journal of Child and Family Studies*, 28(1), pp. 30-41. Available at: https://doi.org/10.1007/s10826-018-1257-3.

Kline, G. H., Wood, L. F. and Moore, S. (2003) Validation of modified family and interparental conflict scales for use with young adults from divorced and non-divorced families, *Journal of Divorce & Remarriage*, 39(3-4), pp. 125-142. Available at: https://doi.org/10.1300/J087v39n03_07.

Liddle, I. and Carter, G. F. (2015) Emotional and psychological well-being in children: the development and validation of the Stirling Children's Well-being Scale, *Educational Psychology in Practice*, 31(2), pp. 174-185. Available at: https://doi.org/10.1080/02667363.2015.1008409.

Lüdemann, J., Rabung, S. and Andreas, S. (2021) Systematic review on mentalization as key factor in psychotherapy, *International Journal of Environmental Research and Public Health*, 18(17), p. 9161. Available at: https://doi.org/10.3390/ijerph18179161.

Luyten, P. et al. (2024) The role of mentalizing in psychological interventions in adults: systematic review and recommendations for future research, *Clinical Psychology Review*, 108, p. 102380. Available at: https://doi.org/10.1016/j.cpr.2024.102380.

Luyten, P., Mayes, L. C., Sadler, L., Fonagy, P., Nicholls, S., Crowley, M. and Slade, A. (2009) *The Parental Reflective Functioning Questionnaire-1 (PRFQ-1)*. Unpublished manual, University of Leuven, Leuven, Belgium.

Luyten, P., Mayes, L. C., Nijssens, L. and Fonagy, P. (2017) The parental reflective functioning questionnaire: development and preliminary validation, *PLoS One*, 12(5), e0176218. Available at: https://doi.org/10.1371/journal.pone.0176218.

Malda-Castillo, J., Browne, C. and Perez-Algorta, G. (2018) Mentalization-based treatment and its evidence-base status: a systematic literature review, *Psychology and Psychotherapy: Theory, Research and Practice*, 92(4), pp. 465-498. Available at: https://doi.org/10.1111/papt.12195.

Masarik, A. S. and Conger, R. D. (2017) Stress and child development: a review of the Family Stress Model, *Current Opinion in Psychology*, 13, pp. 85-90. Available at: https://doi.org/10.1016/j.copsyc.2016.05.008.

Millings, A. (2024) *Support for inter-parental conflict.* ISRCTN Registry. Available at: https://doi.org/10.1186/ISRCTN10266960 (Accessed 31 May 2024).

Millings, A., Hirst, S. L., Sirois, F. and Houlston, C. (2020) Emotional adaptation to relationship dissolution in parents and non-parents: A new conceptual model and measure, *PLoS One*, 15(10), e0239712. Available at: https://doi.org/10.1371/journal.pone.0239712.

MinimPy 0.3 Reference Manual (2011) Available at: https://sourceforge.net/projects/minimpy/files/0.3/.

Nederlof, A. F., Hovens, J. E., Muris, P. and Novaco, R. W. (2009) Psychometric evaluation of a Dutch version of the dimensions of anger reactions, *Psychological Reports*, 105(2), pp. 585-592. Available at: https://doi.org/10.2466/PR0.105.2.585-592.

Owen, J. and Rhoades, G. K. (2012) Reducing interparental conflict among parents in contentious child custody disputes: an initial investigation of the Working Together Program, *Journal of Marital and Family Therapy*, 38(3), pp. 542-555. Available at: https://doi.org/10.1111/j.1752-0606.2010.00215.x.

Porter, B. and O'Leary, K. D. (1980) Marital discord and childhood behavior problems, *Journal of Abnormal Child Psychology*, 8(3), pp. 287-295. Available at: https://doi.org/10.1007/BF00916376.

Saghaei, M. and Saghaei, S. (2011) Implementation of an open-source customizable minimization program for allocation of patients to parallel groups in clinical trials, *Journal of Biomedical Science and Engineering*, 4, pp. 734-739. Available at: https://doi.org/10.4236/jbise.2011.411090.

Scott, N. W., McPherson, G. C., Ramsay, C. R. and Campbell, M. K. (2002) The method of minimization for allocation to clinical trials: a review, *Controlled Clinical Trials*, 23(6), pp. 662-674. Available at: https://doi.org/10.1016/s0197-2456(02)00242-8.

Shifflett, K. and Cummings, E. M. (1999) A program for educating parents about the effects of divorce and conflict on children: an initial evaluation, *Family Relations*, 48(1), pp. 79-89. Available at: https://doi.org/10.2307/585685.

Spybrook, J., Kelcey, B. and Dong, N. (2016) Power for detecting treatment by moderator effects in two- and three-level cluster randomized trials, *Journal of Educational and Behavioral Statistics*, 41(6), pp. 605-627. Available at: https://doi.org/10.3102/1076998616655442.

Tavistock Relationships (no date) *Mentalization Based Therapy for Parenting under Pressure (MBT-PP): A treatment manual for parents under pressure*. Unpublished manuscript. London, UK: Tavistock Centre for Couple Relationships.

van Dijk, R., van der Valk, I. E., Deković, M. and Branje, S. (2020) A meta-analysis on interparental conflict, parenting, and child adjustment in divorced families: examining mediation using meta-analytic structural equation models, *Clinical Psychology Review*, 79, 101861. Available at: https://doi.org/10.1016/j.cpr.2020.101861.

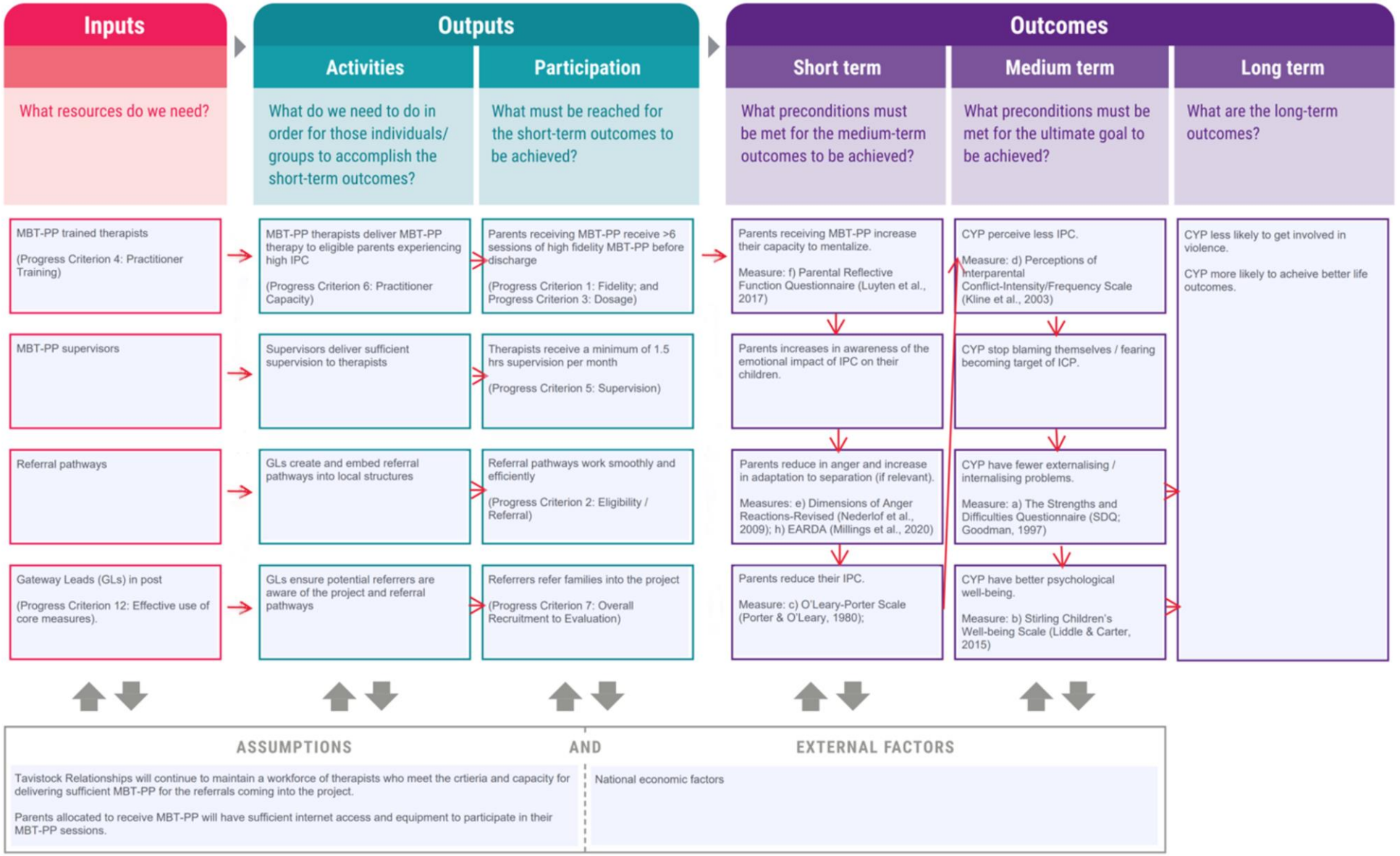
van Eldik, W. M., de Haan, A. D., Parry, L. Q., Davies, P. T., Luijk, M. P., Arends, L. R. and Prinzie, P. (2020) The interparental relationship: meta-analytic associations with children's maladjustment and responses to interparental conflict, *Psychological Bulletin*, 146(7), pp. 553-594. Available at: http://dx.doi.org/10.1037/bul0000233.

Vogt, K. S. and Norman, P. (2018) Is mentalization-based therapy effective in treating the symptoms of borderline personality disorder? A systematic review, *Psychology and Psychotherapy: Theory, Research and Practice*, 92(4), pp. 441-464. Available at: https://doi.org/10.1111/papt.12194.

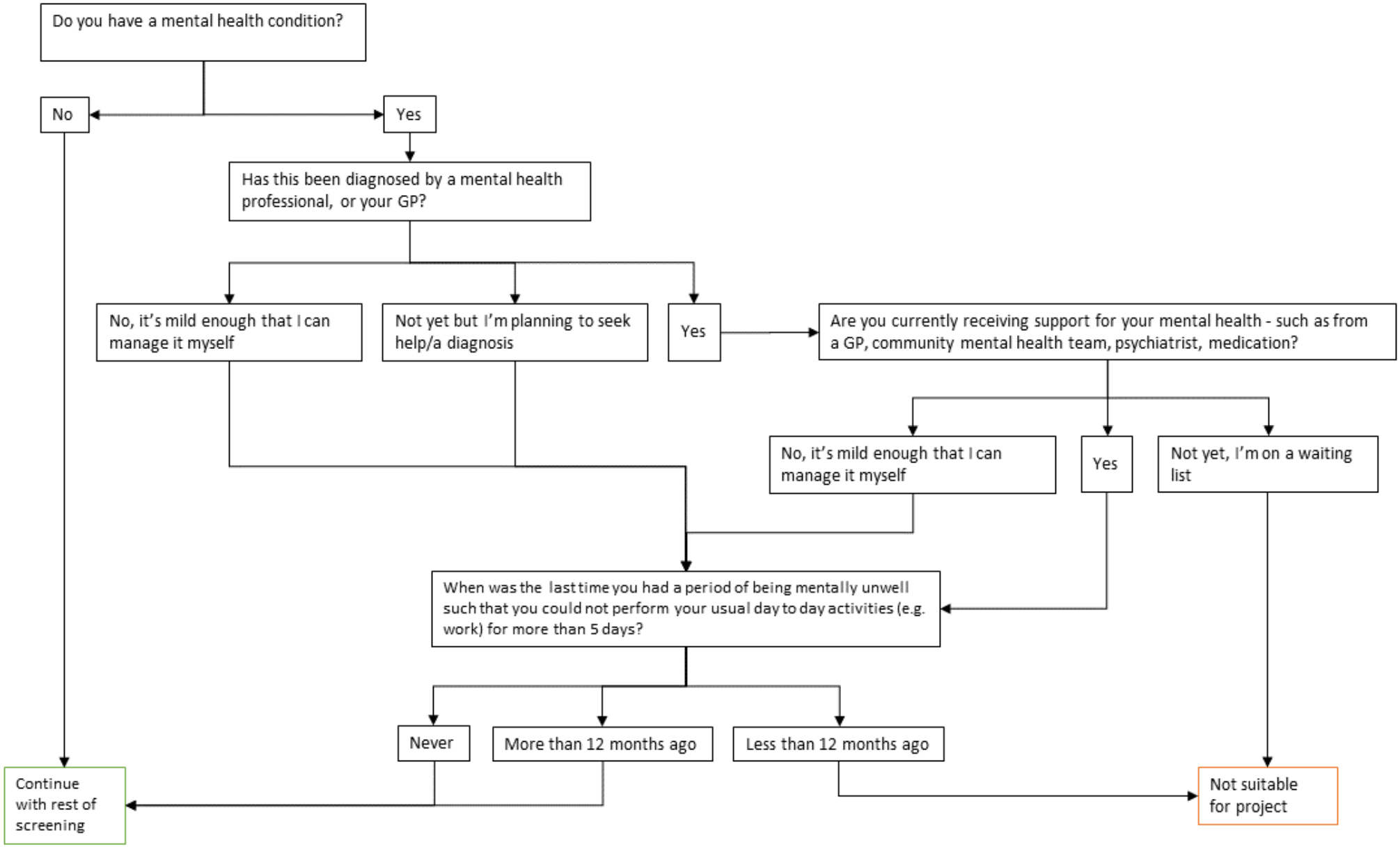
Youth Endowment Fund (2022) *YEF evaluations guidance: pilot studies*. Available at: https://youthendowmentfund.org.uk/wp-content/uploads/2022/03/3.-YEF-Evaluations-Guidance-Pilot-studies-March-2022.pdf (Accessed 10 October 2024).

Appendices

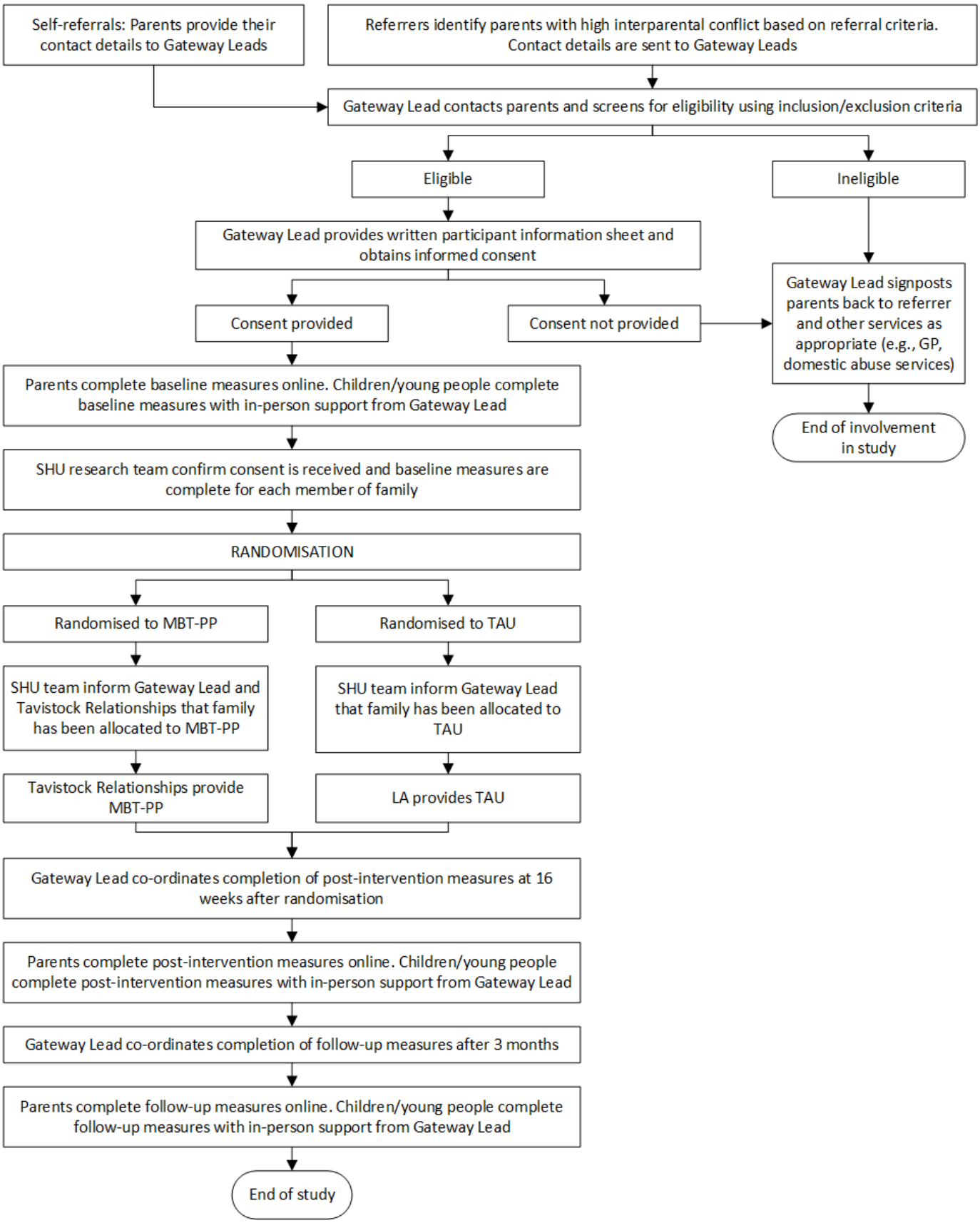
Appendix A The logic model for MBT-PP



Appendix B Flow diagram to support GLs in operationalising mental health exclusion criteria



Appendix C Overview of trial processes



Appendix D Participant information sheet for parents

AN EVALUATION OF MENTALIZATION BASED THERAPY FOR PARENTING UNDER PRESSURE

1. What is this study about?

Conflict between parents can have negative effects on families. The aim of this study is to evaluate a therapy for parents called Mentalization Based Therapy for Parenting under Pressure (MBT-PP). We want to know if taking part in MBT-PP helps parents with their co-parenting relationship, and whether this helps their children's emotions, behaviour and relationships, compared to other services that are available. Our findings will help us understand how best to support families in the future.

The Youth Endowment Fund (YEF), which funds this study, is dedicated to preventing children and young people becoming involved in crime and violence. Once we have finished our study, YEF-approved researchers will explore whether MBT-PP, and other programmes funded by YEF, had an impact over a longer period of time, including whether they reduced young people's involvement in crime and violence. This is explained in more detail below.

2. Who is organising this study?

This study is being organised by Sheffield Hallam University (SHU; www.shu.ac.uk). Our partners for this study are Tavistock Relationships (the organisation providing therapy for parents) and three local authorities (councils): Bristol; Dorset; and Bournemouth, Christchurch & Poole (BCP).

3. Who has reviewed this study?

This study has been reviewed and approved by SHU's Research Ethics Committee, ID = ER50582599.

4. Why have you asked me to take part?

You have been invited to take part because:

- You are experiencing a lot of conflict in your co-parenting relationship (e.g., arguments happen often, are intense, or don't get resolved well). Because of this, you and your child's other parent might benefit from some help to reduce the conflict.

- You have a child/children aged 8-14.

5. Do I have to take part?

No. If you don't want to take part in the study, you don't have to.

We would like as many families as possible to take part, to aid our understanding about what works for young people and their families.

If you choose not to take part in the study, all the usual services provided by your local authority will continue to be available to you. However, the opportunity to have MBT-PP will not be available to you.

6. What happens if I take part?

First, we will ask you to fill in a questionnaire online. This will ask about the relationship you have with your child's other parent, your feelings, and your child/children's emotions, behaviour and relationships. We will also ask you to provide some demographic information and details about your child/children. We expect that this questionnaire will take you around 20-30 minutes.

We will also ask your child/children who are aged 8-14 years to fill in a questionnaire that asks about their wellbeing and perceptions of family relationships. We will arrange to help your child/children fill in these questionnaires at a suitable time. We expect that this will take them up to 20 minutes.

Next, we will randomly allocate your family to one of two groups: one group will have MBT-PP and the other group will be supported by their local authority. Because allocation will be random, we can't tell you now whether you will be offered MBT-PP or support from your local authority. However, we believe that both MBT-PP and the local authority support should be helpful for reducing conflict.

Parents in the families who have MBT-PP will be offered ten sessions with a therapist. Sessions will be conducted remotely using a video-conferencing platform e.g., Zoom or Teams. There will be an initial session with each parent individually. The remaining sessions will be with both parents and the therapist, during which parents will discuss their co-parenting relationship and learn new ways of responding during disagreements. The therapist will record the sessions so that we can check what has been done in the sessions.

Parents in the families who are supported by their local authority will receive the usual support for families in conflict that is available in their local area. This may include accessing online resources for improving relationships and/or support from a worker from the local authority or from an external service commissioned by the local authority.

After around four months, we'll ask you to fill in another online questionnaire about the relationship you have with your child's other parent, your feelings, and your child's emotions, behaviour and relationships, so that we can see if anything has changed.

If you still need support to reduce conflict at this time, all the usual services provided by your local authority will be available to you.

We'll ask you to fill in a final questionnaire three months after this (i.e., three online questionnaires in total).

We will also ask your child/children to fill in the same questionnaire as before at these time points (four months after beginning the study, then three months after that), and we will arrange to help them do this at a suitable time.

7. What are the benefits or advantages of taking part?

As thanks for your time spent filling in the questionnaires, we will give you a £10 shopping voucher for each questionnaire you complete, plus a bonus £10 voucher for completing all three questionnaires. This will also apply for your child's other parent and for your child/children, meaning that a family with one child completing all three questionnaires will receive vouchers to the value of £120.

Taking part in this study means that you have the chance to receive MBT-PP free of charge, which is not otherwise available in your local area.

8. What are the risks or disadvantages of taking part?

Taking part in the study will take up some of your time, both from filling in the questionnaires and when attending sessions with your therapist (if you are allocated to MBT-PP) or when receiving support from your local authority (if you are allocated to local authority support).

Having therapy or support to deal with conflict in a relationship can be difficult in the short-term, as it may involve thinking about things that are upsetting for you. However, addressing these problems is expected to be beneficial for you and your family in the long-term.

Occasionally, someone may feel upset about a question that is asked during the study. If you or your child feel upset by any of the questions asked as part of this study and want to speak to someone about it,

you/they can speak to the person who provided this information sheet, or the person who is supporting you with your conflict (your MBT-PP therapist or staff from your local authority / partner organisations). You can also contact the Project Lead, Prof. Abigail Millings, using the contact details below.

If, at any point, you or your child feel that you need further help with mental health or wellbeing, please contact your GP, who may be able to refer you on for further support. You can also contact an external support service such as The Samaritans (Tel. 116 123, www.samaritans.org) or Childline (Tel. 0800 1111, www.childline.org.uk).

9. What if I change my mind?

If you want to stop taking part in the study, you are free to do so at any point. Contact the Project Lead (details below) if you decide to stop participating. You will still be able to receive support from your local authority or your MBT-PP therapist if you have been allocated to one. All the usual services provided by your local authority will continue to be available to you.

If you decide to withdraw, you do not have to give a reason but you should tell us as soon as possible. It will not be possible to delete any personal information already collected because we will be using this information, along with all of the information we have gathered from the other participants, to carry out our evaluation and to write our report.

Once information goes into the YEF archive it cannot be deleted as that would affect the quality of the archived data for use in future research.

10. If I have questions about taking part, who do I speak to?

You can contact the Project Lead, Prof. Abigail Millings, A.Millings@shu.ac.uk, 0114 225 2612.

11. What happens to the information you collect?

During this study

- All of the information that you and your child/children give us will be stored securely.
- We will treat the information that you and your child/children share with us as confidential, but we may have to break confidentiality if you or they tell us something that makes us concerned about someone being at risk of harm. If this happens then we will usually discuss the issue with you/them first. You can find more information in our Safeguarding Policy (https://www.shu.ac.uk/about-us/governance/governance-documents/safeguarding).
- Selected audio-recordings of MBT-PP sessions will be shared by Tavistock Relationships with the research team at Sheffield Hallam University. Recordings will be shared using a secure data transfer service.
- We will write a report about our findings based on the information we receive from all participants. This report will not contain any personal information about the people who took part in the study and it will not be possible to identify individuals from the report. The report will be published on the YEF's website. We might also use the findings in academic contexts (e.g., teaching, research articles, and conference presentations).

After the study

- Once we have finished our study, we will share information about the children of people who have taken part in the study with the Department for Education (DfE). The DfE will replace all identifying information about these children (their name, date of birth, home address) with the children's unique Pupil Matching Reference number in the DfE's National Pupil Database. Once this has been

done, it is no longer possible to identify any individual child from the study data. This process is called pseudonymisation.

- The DfE will transfer the pseudonmyised information to the YEF archive, which is stored in the Office for National Statistics' Secure Research Service. Information in the YEF archive can only be used by approved researchers to explore whether MBT-PP had an impact over a longer period of time for example, whether being part of a project reduces a child's likelihood of being excluded from school or becoming involved in criminal activity. Using the unique Pupil Matching Reference numbers added to the data by the Department for Education, it will be possible to link the records held in the YEF archive to other public datasets such as education and criminal justice datasets.
- The YEF archive is protected by the Office for National Statistics' 'Five Safes' framework. The information can only be accessed by approved researchers in secure settings and there are strict restrictions about how the information can be used. Information in the YEF archive cannot be used by law enforcement bodies or by the Home Office for immigration enforcement purposes. You can find more information about the YEF archive and the Five Safes in the guide for participants on the YEF's website: https://youthendowmentfund.org.uk/wp-content/uploads/2021/07/YEF-Data-Guidance-Participants.pdf. We encourage all parents to read the YEF's guidance for participants before deciding to take part in this study.
- SHU will keep your child's/children's personal information for 12 weeks after we have transferred the data to DfE for archiving, which we expect to be around September 2025. The YEF will keep information in the YEF archive for as long as it is needed for research purposes. The YEF we will carry out a review every five years to assess whether there is a continued benefit to storing the information in the archive, based on its potential use in future research.
- SHU will remove any information that could identify individuals and securely store the fully anonymised data from this study for a period of at least 10 years from completion. Audio-recordings of MBT-PP therapy sessions will be deleted at the end of the study.

Equality and Diversity

We recognise that people experience hostility, prejudice, and discrimination based on their cultural background, marginalised identities, and beliefs. We are committed to inclusivity and belonging, and we want to ensure that participation in this research is a safe and friendly experience for people of all cultural backgrounds. We also recognise that this is not a straightforward endeavour, but requires continuous improvement, so we are grateful for any feedback you may have for us to help us improve. Feel free to contact our Research Inclusion Lead:

Helen Birtwhistle
Centre for Behavioural Science and Applied Psychology (CeBSAP)
Sheffield Hallam University
Heart of the Campus
Collegiate Crescent
Sheffield
S10 2BQ
Email address: cebsap@shu.ac.uk Telephone number: 0114 2255046

Legal basis for research for studies

The University undertakes research as part of its function for the community under its legal status. Data protection allows us to use personal data for research with appropriate safeguards in place under the legal

basis of public tasks that are in the public interest. A full statement of your rights can be found at https://www.shu.ac.uk/about-this-website/privacy-policy/privacy-notices/privacy-notice-for-research. Sheffield Hallam University are the data controllers during the study.

You should contact the Data Protection Officer, Helen Williamson, at DPO@shu.ac.uk if you have a query about how your data is used by the University, if you would like to report a data security breach (e.g., if you think your personal data has been lost or disclosed inappropriately), or if you would like to complain about how the University has used your personal data.

Once information is transferred to the DfE to be pseudonymised, Sheffield Hallam University hand over control for protecting your personal information to the YEF. The YEF is the controller of the information in the YEF archive and use personal data under the legal basis of performing tasks in the public interest.

For data queries after this time, please contact the YEF at hello@youthendowmentfund.org.uk. Further information is available in YEF's guidance for participants, available at https://youthendowmentfund.org.uk/evaluation-data-archive/.

We always encourage you to speak to us first, but if you remain unsatisfied you also have the right to make a complaint at any time to the Information Commissioner's Office (ICO), the UK supervisory authority for data protection issues: https://ico.org.uk/make-a-complaint/.

You should contact the Head of Research Ethics if you have concerns with how the research was undertaken or how you were treated: Dr Mayur Ranchordas, Sheffield Hallam University, Howard Street, Sheffield S1 1WBT. Telephone: 0114 225 5555 Email: m.ranchordas@shu.ac.uk

Appendix E Participant information sheet for children and young people

AN EVALUATION OF MENTALIZATION BASED THERAPY FOR PARENTING UNDER PRESSURE

1. What is this study about?

Sheffield Hallam University is doing a study of people who are having a talking therapy called Mentalization Based Therapy for Parenting under Pressure (MBT-PP).

We are trying to find out whether this talking therapy for parents helps them to get along better and whether it helps their children to feel better too. Our findings will help us understand how best to support families in the future.

The Youth Endowment Fund, or YEF for short, is giving us money to do this study. Once we have finished our study, the information we collect will be stored by the YEF and used to see how MBT-PP, and other programmes funded by YEF, affected families over a longer period of time. This is explained in more detail below.

2. Why have you asked me to take part?

You have been invited to take part because:

- Your parents are taking part in the study.

- You are aged between 8 and 14 years old.

3. Do I have to take part?

No. If you don't want to take part in the study, you don't have to. If you don't want to take part, tell your parent or guardian or the Project Lead (details below).

We would like as many families as possible to take part, to help us understand what works for young people and their families.

Even if you don't take part in the study, your parents can still take part.

4. What happens if I take part?

We will ask you to answer some questions (fill in a questionnaire), with help from an adult involved in this project. These questions will ask about your recent thoughts and feelings, and what you think and feel about the way your parents get along. We expect that this questionnaire will take you around 20 minutes to fill in and we will arrange to complete these with you at a time that suits you and your family.

We will ask you to answer the same questions again after around four months, and then again three months after that, so that we can see whether anything has changed. This means we will ask you to fill in the questionnaire at three different times.

5. What are the benefits of taking part?

As thanks for your time, we will give you a £10 shopping voucher for each questionnaire you complete, plus a bonus £10 voucher for completing all three questionnaires. This means that if you complete all three questionnaires, you will receive vouchers worth £40.

6. What are the risks or disadvantages of taking part?

Taking part in the study will take up some of your time.

Occasionally, someone may feel upset about a question that is asked during the study. If you feel upset by any of the questions asked as part of this study, you should tell your parent or guardian. If you want to talk to someone else about it, you can talk to the person who provided this information sheet, or talk to a trusted adult at your school (e.g., your class teacher, learning mentor, head of year). You can also contact the Project Lead, Dr Abigail Millings, using the contact details below.

If, at any point, you feel that you need further help with how you are feeling, you can see your family doctor (your GP), who may be able to refer you on for further support. You can also contact Childline (Tel. 0800 1111, www.childline.org.uk). Calls to Childline are free from landlines and mobiles in the UK, and they won't show up on your phone bill.

7. What if I change my mind?

If you want to stop taking part in the study, you are free to do so at any point. Contact the Project Lead (details below), or ask an adult to do this for you, if you decide to stop participating.

If you decide to stop taking part, you don't have to give a reason but please tell us as soon as possible. We won't be able to delete any information that we've already collected from you because we will be using this information, along with the information we have gathered from other people taking part in the study, to carry out our work.

8. Who is organising this study?

This study is being organised by Sheffield Hallam University (www.shu.ac.uk). This study has been reviewed and approved by Sheffield Hallam University's Research Ethics Committee, ID = ER50582599.

9. What happens to the information you collect?

During this study

- We always keep your information safe. During the study, we only let our research team look at your information.
- We will keep what you tell us private unless we think that you or someone else might be at risk of harm. If this happens, then we will usually talk to you first to tell you why we want to talk to another person or organisation.
- We will write a report about what we find, but the report won't include your name or any other information that could be used to identify you. The report will go on the YEF's website and anyone will be able to read it. We might also use the findings in articles that we write, on our website, or in presentations.

After the study

- When we finish the study, we'll give your information to the YEF and they will become the 'controller' of it. They will keep your information in a safe place called the YEF archive. You can find more information about the YEF archive on the YEF's website: https://youthendowmentfund.org.uk/wp-content/uploads/2021/07/YEF-Data-Guidance-Participants.pdf.
- Before your information goes into the YEF archive, the Department for Education will take out your name and other personal details like your address. This means that no one who looks at the information in the YEF archive will know who you are.
- In the future, people can ask to use the YEF archive to do more studies to find out whether MBT-PP, and other projects like ours, have helped young people. Only researchers who are approved by the YEF will be able to look at the archive. The police can't use the information in the YEF archive.

- Information will be kept safely in the YEF archive for as long as it is needed for future research.
- Sheffield Hallam University will keep your personal information for 12 weeks after we give information from the study to the YEF. After this we will take out your name and other personal details from the information we keep so no one will be able to know who you are from the data.

10. If I have questions about taking part, who do I speak to?

You can contact the Project Lead, Dr Abigail Millings, A.Millings@shu.ac.uk, 0114 225 2612.

How we comply with the law

We will only use your information if the law says it's ok. Because this study is interesting and important to lots of people, the law says we can use your information to do this kind of work.

The law gives you rights over how we can use your information. You can find full details of these rights at https://www.shu.ac.uk/about-this-website/privacy-policy/privacy-notices/privacy-notice-for-research. Sheffield Hallam University are the data controllers during the study.

If you have any questions about how we use your information, or if you want to complain, you can contact our Data Protection Officer, Helen Williamson, at DPO@shu.ac.uk.

Once information is transferred to the DfE, Sheffield Hallam University hand over control for protecting your personal information to the YEF. The YEF is the controller of the information in the YEF archive.

For data queries after this time, please contact the YEF at hello@youthendowmentfund.org.uk. Further information is available in YEF's guidance for participants, available at https://youthendowmentfund.org.uk/evaluation-data-archive/.

We always encourage you to speak to us first, but you also have the right to make a complaint to the Information Commissioner's Office (ICO). You can find more information about the ICO and how to make complaint to them on their website: https://ico.org.uk/make-a-complaint/.

You should contact the Head of Research Ethics if you have concerns with how the research was undertaken or how you were treated: Dr Mayur Ranchordas, Sheffield Hallam University, Howard Street, Sheffield S1 1WBT. Telephone: 0114 225 5555 Email: m.ranchordas@shu.ac.uk

Appendix F Fidelity checklist MBT-PP

Session identifier:

Site:

Coder:

Observed interactions

To what extent did you observe the following during the session?

1. Therapist **asks clients how they are feeling/thinking** without making assumptions

Not observed	Possibly observed	Clearly observed
0	1	2

2. Therapist **voices their observation of behaviour and/or feeling** between parents

Not observed	Possibly observed	Clearly observed
0	1	2

3. Therapist **voices their observation of repeated patterns of behaviour** that impact on parents' feelings

Not applicable	Not observed	Possibly observed	Clearly observed
☐	0	1	2

4. Therapist **checks to see if parents recognise/can see the behaviours described** by them, to check that they are correct

Not applicable	Not observed	Possibly observed	Clearly observed
☐	0	1	2

5. Therapist encourages parents talk **to each other directly within the session** to voice their thoughts and feelings

Not observed	Possibly observed	Clearly observed
0	1	2

6. Therapist **pauses an interaction between parents in the session** to ask what person X is feeling

Not applicable	Not observed	Possibly observed	Clearly observed
☐	0	1	2

7. Therapist **moves from discussing a specific interaction** that occurred within the session **to make more general observations about patterns** and associated feelings

Not applicable	Not observed	Possibly observed	Clearly observed
☐	0	1	2

8. Therapist works with parents to **agree a name for an unhelpful pattern of interaction**

Not applicable	Not observed	Possibly observed	Clearly observed
☐	0	1	2

9. Therapist encourages each parent to **talk about their own and the other parent's constructive solutions** to challenging patterns

Not applicable	Not observed	Possibly observed	Clearly observed
☐	0	1	2

10. Therapist encourages parents to **reach an agreement on how to avoid negative patterns of interaction and/or to regulate feeling states**

Not applicable	Not observed	Possibly observed	Clearly observed
☐	0	1	2

11. Therapist **checks feelings and understanding at the end of the session**

Not observed	Possibly observed	Clearly observed
0	1	2

Core skills

To what extent did the therapist demonstrate the following skills?

12. Show warmth and respect for each parent

Not demonstrated	Possibly demonstrated	Clearly demonstrated
0	1	2

13. Listen and give both parents a sense of being understood – checking out understanding regularly

Not demonstrated	Possibly demonstrated	Clearly demonstrated
0	1	2

14. Being inclusive and even-handed with both parents

Not demonstrated	Possibly demonstrated	Clearly demonstrated
0	1	2

15. Identifying and highlighting strengths, particularly in mentalising

Not demonstrated	Possibly demonstrated	Clearly demonstrated
0	1	2

16. Focus on strengths between parents alongside working on the difficulties parents are struggling with

Not demonstrated	Possibly demonstrated	Clearly demonstrated
0	1	2

17. Intervene and manage non-mentalising interactions in the session if they occur

Not applicable	Not demonstrated	Possibly demonstrated	Clearly demonstrated
☐	0	1	2

18. Refocus sessions if they wander from the core task

Not applicable	Not demonstrated	Possibly demonstrated	Clearly demonstrated
☐	0	1	2

19. Maintain interest in the couple state of mind at all times to notice and re-establish it if you find it goes 'offline'

Not applicable	Not demonstrated	Possibly demonstrated	Clearly demonstrated
☐	0	1	2

20. Provide clear time boundaries

Not demonstrated	Possibly demonstrated	Clearly demonstrated
0	1	2

21. Speak with confidence about MBT-PP and its aims (where applicable – likely more relevant in earlier sessions).

Not applicable	Not demonstrated	Possibly demonstrated	Clearly demonstrated
☐	0	1	2