EVALUATION PROTOCOL

Becoming a Man, a two-armed randomised controlled trial.

The Ending Youth Violence Lab

Principal investigator(s): Tom McBride

YOUTH ENDOWMENT FUND

Becoming a Man (BAM)

Evaluation protocol

Evaluating institution: Ending Youth Violence Lab
Principal investigator(s): Tom McBride

Project title[1]	Becoming a Man, a two-armed randomised controlled trial.
Developer (Institution)	London Borough of Islington (delivery partner in this study) Mental Health Foundation (historic delivery partner) Youth Guidance (US Developer)
Evaluator (Institution)	The Ending Youth Violence Lab
Principal investigator(s)	Tom McBride
Protocol author(s)	Tom McBride, Lilli Wagstaff, Lisa Mueller, Jack Martin
Trial design	Two-armed randomised controlled trial with individual level random allocation within schools
Trial type	Efficacy study with an internal pilot and integrated implementation and process evaluation
Evaluation setting	Mainstream Secondary Schools across Greater London
Target group	Boys in years 8-10. It is expected that the vast majority of boys will present with lower

[1] Please make sure the title matches that in the header and that it is identified as a randomised trial as per the CONSORT requirements (CONSORT 1a).

	secondary, upper secondary level, or a tertiary level of need.
Number of participants	1,080 children from 9 schools
Primary outcome and data source	Behavioural difficulties (Self-report - Strengths and Difficulties Questionnaire - Total Score) only for those assessed as being at upper-secondary or tertiary level of need
Secondary outcome and data source	Offending (a binary indicator of whether a boy has been arrested for any offence - violent and non-violent - in the 18 months following randomisation, using local police data) only for those assessed as being at upper-secondary or tertiary level of need[2] Behavioural difficulties (Self-report - Strengths and Difficulties Questionnaire - Total Score) for all boys. Self-esteem (Self-report - Rosenberg Self-Esteem Scale). Self-control and emotional regulation (Self-report - Difficulties in Emotion Regulation Scale). Social skills (Self-report - Strengths and Difficulties Questionnaire - prosocial score) Externalising behaviour (Self-report - Strengths and Difficulties Questionnaire - externalising problems score) School attendance (school data)

[2] Note that data will be gathered on offending for everyone within the trial and report this as an exploratory analysis. However, since those assessed as being primary or lower-secondary level of need are likely to be some distance from offending they will not be included in the secondary analysis..

	Exclusions, suspensions and transfers (school data)

Protocol version history

Version	Date	Reason for revision
1.0	20/10/2026	

Any changes to the design or methods need to be discussed with the YEF Evaluation Manager and the developer team prior to any change(s) being finalised. Describe in the table above any agreed changes made to the evaluation design. Please ensure that these changes are also reflected in the SAP (CONSORT 3b, 6b).

Table of contents

Protocol version history **3**
Table of contents **4**
Study rationale and background **6**
Intervention **10**
Programme content and adaptation 12
BAM Theory of Change 16
Research questions or study objectives 18
Pilot 18
Efficacy Trial 19
Progression criteria **37**
Impact evaluation: Design **43**
Table 4: Trial design **44**
Randomisation 46
Participants 48
Process 48
Referral 50
Inclusion and exclusion criteria for young people 50
Inclusion and exclusion criteria for schools 52
Sample size calculations 53
Impact Evaluation: Methods and data collection **56**
Data sources 56
Baseline measures 58
Outcome measures 59
Primary Outcome 59
Secondary Outcomes 59
Exploratory outcomes 60
Analysis 68
Implementation and process evaluation **70**
Research questions 70
Research methods 72
Administrative/monitoring data 72
Surveys 72
Qualitative research 73
Fidelity 75
1. Self-Reported Adherence (Implementation Checklists) 76
2. Expert Observation (Replication Specialist Assessment) 76
3. Curriculum Progression (Dosage/Reach) 76
4. Fidelity Targets 77
5. School Enabling Context 77
Analysis 78
Data analysis 78

Cost data reporting and collecting **78**
Diversity, equity and inclusion **79**
Ethics and registration **84**
Data protection **84**
Stakeholders and interests **86**
Risks **88**
Timeline **90**
Appendix 1 - Adaptation Approach **91**
Appendix 2: Theory of Change **96**
Appendix 3: Case Studies **103**
Appendix 4: Referral / Recruitment Rate Table **104**
Appendix 5: BAM Contamination Risk: RAG Rating **105**
BAM Contamination Risk: RAG Rating **105**
Rating Definitions 105
RAG Ratings 105

Study rationale and background

Violent and aggressive behaviours among young people present a significant societal challenge, often leading to school exclusion, limiting opportunities, and long-term involvement in the criminal justice system. These behaviours are frequently rooted in complex systemic issues, including developmental adversity, poverty, discrimination, and trauma, where children often develop "automatic" defensive behaviours as adaptive survival mechanisms[3]. To address this, it is important to develop and rigorously evaluate interventions that support young people to navigate high-risk environments by fostering social-emotional skills and resilience within their school community[4]. Additionally, the government's current strategy aims to steer boys away from misogynistic influences and provides dedicated funding for targeted behavioural initiatives for at-risk boys[5], while the RSHE guidance requires a curriculum that fosters emotional expression and healthy identity[6]. Together, these policies could help transform optional psycho-social programmes into essential, policy-aligned requirements for educational institutions.

School-based Social and Emotional Learning (SEL) programmes are frequently cited as a promising approach to fostering social-emotional skills[7]. SEL interventions focus on fostering skills such as self-awareness, decision-making, relationship-building, and communication. There is substantial evidence that effective SEL education can improve school bonding and academic performance while reducing conduct problems and aggressive behaviour. There is also some promising evidence that targeted social and emotional learning may reduce violence by 43% and crime and offending by 31%[8]. However, the evidence base for these findings is still emerging and there is a lack of high quality studies[9].

By embedding SEL interventions within schools, programmes can potentially "slow down" high-stakes decision-making processes, helping young people navigate high-stress environments without resorting to violence.

It is also crucial for interventions to address disparities in how the complex systems impact individuals from minoritised backgrounds. Recent evidence highlights that while Black children aged 10-17 comprise approximately 4% of the general population, they account for 15% of arrests and 29% of the youth custody population[10]. Furthermore, in 2023/24, 34% of

[3] Heller et al. (2017)

[4] https://youthendowmentfund.org.uk/wp-content/uploads/2024/07/YEF.-BAM-Pilot-Report.-July-2024-1.pdf

[5] https://www.gov.uk/government/publications/freedom-from-violence-and-abuse-a-cross-government-strategy

[6] https://www.gov.uk/government/publications/relationships-education-relationships-and-sex-education-rse-and-health-education

[7] https://youthendowmentfund.org.uk/toolkit/social-skills-training/

[8] https://youthendowmentfund.org.uk/toolkit/social-and-emotional-learning-targeted/

[9] Ibid.

[10] https://www.gov.uk/government/news/annual-statistics-a-system-failing-black-children

homicide victims aged 16-24 were Black[11], a figure more than six times their share of the population. Children from Black and Mixed ethnic groups accounted for 13% and 8% of all prosecutions, respectively compared to 3% for White defendants[12]. Young people who are eligible for free school meals, have special educational needs and are suspended or excluded from school show a higher incidence of cautions and sentencing for an offence compared to the wider pupil population[13]. Young people from minority ethnic groups are also disproportionately represented in the prison system: 53% of the prison population of under 18-year olds in England and Wales in 2020 were from minority ethnic groups (32% Black, 12% Mixed, 8% Asian, 1% Other)[14]. These racial disparities are also reflected in school exclusions: in the 2023/24 academic year Gypsy, Roma and Traveller pupil populations had the highest numbers of suspension and permanent exclusions in England, whilst mixed white and black Caribbean pupils had the second highest rates[15]. Addressing these disparities requires interventions that are not only evidence-based but also culturally responsive and capable of engaging young people from minoritised communities who are disproportionately affected by violence.

To fill the gap in robust UK-based evidence for such interventions, the Youth Endowment Fund (YEF) identified 'Becoming a Man' (BAM) as a promising programme for evaluation. BAM is a targeted SEL intervention originally developed in Chicago, designed specifically to help children engage in responsible decision-making, ultimately improving school engagement and reducing interactions with the criminal justice system. Practitioners – known as BAM counsellors – deliver the intervention in schools with groups (called BAM Circles) of 8–12 participants, called BAM scholars, over two school years[16] - although for this evaluation the programme is being adapted to be delivered in a single school year, see below for more details. The programme also has a strong track record with minority ethnic groups in the US.

Since its UK launch in 2020, the BAM programme has evolved through a process of cultural adaptation that centres a strengths-based, non-pathologising ethos. MHF recognised that

[11] https://youthendowmentfund.org.uk/wp-content/uploads/2025/07/Beyond_the_Headlines_2025.pdf

[12] Statistics on Ethnicity and the Criminal Justice System, 2024 (HTML). (2025, November 27). GOV.UK. https://www.gov.uk/government/statistics/ethnicity-and-the-criminal-justice-system-2024/statistics-on-ethnicity-and-the-criminal-justice-system-2024-html#:~:text=Age%20groups%20%5Bfootnote%2025%5D%20*%20This%20decrease,for%20Black%20individuals%2C%20which%20decreased%20by%203%25.

[13] Ibid.

[14] Ethnicity and the Criminal Justice System, 2020. (2021, December 2). GOV.UK. https://www.gov.uk/government/statistics/ethnicity-and-the-criminal-justice-system-statistics-2020/ethnicity-and-the-criminal-justice-system-2020#:~:text=6.1%20Prison%20Population,50%20or%20over%20were%20White.&text=In%202020%2C%20Black%20offenders%20made,all%20prisoners%20aged%20under%2025.

[15] https://explore-education-statistics.service.gov.uk/find-statistics/suspensions-and-permanent-exclusions-in-england/2024-25-autumn-term

[16] https://youthendowmentfund.org.uk/wp-content/uploads/2023/07/MHF-BAM.-YEF-Feasibility-study.-2023.pdf

young men are not the 'problem' but are individuals navigating a context of systemic violence and structural inequality; therefore, the role of BAM is not to place the burden of system change on young people, but to provide a safe, reflective environment for harm mitigation and personal growth. By viewing BAM Scholars through the lens of their potential rather than their challenges, BAM Counsellors provide the safety and belonging they need to respond to their environments consciously. This work extends beyond the BAM Circle through the full-time embedded Counsellors, who work to disrupt harmful institutional practices - such as harsh discipline and school exclusions - by engaging school staff in professional development and trauma-informed support. Ultimately, BAM seeks to empower young men to navigate the world as it is, while actively working within school systems to prevent the reproduction of further harm.

The extensive and iterative piece of work on cultural adaptation led by MHF has helped to ensure the cultural relevance of the programme to British young people. Part of this adaptation required an analysis of the specific systemic factors that reproduce violence and inequalities in the lives of young people. BAM was thus positioned as a programme that would support young people to navigate these conditions in ways that align better with their goals and aspirations.

Following a feasibility study (2020–2022) and a pilot study (2021–2023) which demonstrated evidence of promise and implementation feasibility, this next phase will involve a larger-scale evaluation to establish the impact of the programme on key outcomes such as behavioural difficulties and offending. Historically, BAM delivery in the UK has been open to young people from all backgrounds and participation has not focussed on serving populations at higher risk of exclusion or criminalisation. It has been run by the MHF in Lambeth and Islington and has not focussed on cohorts with specific vulnerabilities in terms of youth violence.

Study Context

Becoming a Man (BAM) has a strong evidence base in the United States, where it was developed by Youth Guidance. It has been evaluated through two rigorously conducted randomised controlled trials (RCTs) in Chicago, involving predominantly African American and Hispanic young boys from economically deprived communities. Both studies found that BAM reduced arrests and improved school outcomes[17,18].

Despite this success in the US, the programme had not been delivered in the United Kingdom prior to 2020. To address this, the YEF commissioned a staged evaluation to test the transferability of the model to the UK education system.

[17] Heller et al., 2013

[18] Heller et al., 2017

- **Feasibility Study (2020–2022)** To test whether this model could be transported to the UK, a feasibility study was conducted with a first cohort of 95 scholars across two secondary schools and one Pupil Referral Unit (PRU) in Lambeth. This phase focused on adaptation and deliverability. The study concluded that BAM was feasible to deliver in mainstream UK secondary schools, with high attendance rates (98% of scholars attended at least one session) and positive perceptions from participants. However, it also identified that the programme was less feasible in a PRU setting due to the severity of needs and inconsistent attendance, leading to the withdrawal of the programme from the PRU for the subsequent pilot phase.
- **Pilot Study (2021–2023)** Building on the feasibility findings, a pilot study was conducted with a second cohort of 97 scholars across three mainstream secondary schools in Lambeth. The cohort was predominantly Black/Black British (66%). This phase assessed "evidence of promise" and found:
    - **Implementation successes:** The curriculum was successfully adapted to the UK context (e.g., language and cultural references) while retaining core elements.
    - **Positive outcomes:** Qualitative evidence showed students found the programme "fun, engaging, and relevant". Quantitatively, while the sample size was too small for definitive impact statements, a statistically significant positive change was observed in **empathy** scores on the Holistic Student Assessment (HSA).
    - **Key mechanisms:** Interviews suggested the programme successfully engaged scholars in "action and reflection," helping them internalise values like Integrity and Accountability.
    - **Poor dosage:** No group completed all 30 lesson units in the manual. The total number of weekly group sessions a scholar attended was on average 17 sessions in Year 1 and 12 in Year 2. While this met the minimum threshold of 29 total sessions, it was far below the goal of 45 planned sessions.
    - **Group imbalances**: Recruitment failed to maintain the target split across tiers of needs (15%, 70% and 15%), resulting in groups skewed heavily towards high-need scholars (e.g. 56% Tier 3).
    - Unintended risk to attainment: School staff raised concerns that regularly **missing academic classes** for BAM posed a risk to educational attainment, particularly for higher need boys in multiple school-based interventions.
- **Efficacy Trial (2026-2030)** While the pilot study demonstrated that BAM is deliverable and shows promise in a UK setting, it was not designed or powered to attribute impact to the intervention compared to a counterfactual group. Given the robust US evidence and the promising signals from the UK pilot (specifically regarding social-emotional development and empathy), a full efficacy trial is now required. This larger-scale RCT is necessary to provide rigorous, quantitative

estimates of BAM's impact on preventing youth violence and improving educational outcomes in the UK context, ensuring that resources are invested in interventions with proven effectiveness. Thus, the programme will focus on participation from a cohort with specific vulnerabilities for youth violence, contrary to historical BAM delivery in the UK.

This protocol presents the proposed approach to the efficacy study, which will include an internal pilot and an integrated implementation and process evaluation (IPE).

Intervention

Becoming a Man (BAM) is a school-based social and emotional learning (SEL) group intervention originally developed in Chicago by Youth Guidance and historically delivered in the UK by the Mental Health Foundation. Race equity was a particular focus during the development of BAM: The intervention was developed with a particular aim of working with underserved Black and Hispanic boys and young men in Chicago, many of whom were navigating difficult life circumstances. BAM was co-produced with these communities. Throughout the adaptation to the UK context the intervention maintained the focus on race equity through the involvement of Black Thrive - a racial equity activism group - to support the refinement to a new cultural context and the aim of recruiting counselors whose ethnic background aligns with the community they serve. The London Borough of Islington (Islington) is the delivery partner for this current study and their Young Islington Service will be leading the delivery. While The Mental Health Foundation as an organisation will cease all involvement with the programme delivery, all staff involved in delivering BAM at MHF have been brought onto the team at Islington to deliver BAM. Young Islington delivers universal, targeted and specialist services to children and young people in the London borough of Islington. Some of these children and young people require early intervention and prevention services only, whilst some, who have more complex needs, require more intensive support and interventions. Young Islington sits under Islington's Children's Services department and generally works with service users up to the age of 18 (although some of its services support young people up to the age of 25). BAM seeks to improve Social and Emotional Competencies and has been proven to also have an impact on school engagement and reduce interaction with the criminal justice system[19] [20]. It operates on a theory of change that integrates clinical theory—specifically Cognitive Behavioural Therapy (CBT)—with "Men's Work" and rites of passage. The central hypothesis is that many children in high-stress environments develop "automatic" behaviour, or adaptive survival instincts, often resulting in adolescents turning

[19] Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S., and Pollack, H. A. (2017). "Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago." Quarterly Journal of Economics, 132(1), 1-54.

[20] Heller, S., Pollack, H. A., Ander, R., and Ludwig, J. (2013). "Preventing youth violence and dropout: A randomized field experiment." Cambridge MA: National Bureau of Economic Research.

to aggressive or defensive behaviours when confronted with circumstances in which such behaviours could be counterproductive or damaging to themselves and others[21]. As they navigate the transition to manhood, BAM supports them to internalise core values as "positive assets," empowering them to slow down this automatic thinking and make responsible decisions. Through a process of "action and reflection" within a safe, consistent relationship with a pro-social male adult, BAM challenges these automatic responses and supports scholars in internalising six core values: Integrity, Accountability, Self-Determination, Positive Anger Expression, Respect for Womanhood, and Visionary Goal Setting. Through the use of diverse curriculum activity approaches (lectures, discussions, interactive activities, video education, group missions) BAM aims to accommodate different learning and interactivity needs of neurodiverse students.

To implement this model effectively, the programme targets adolescent boys, typically aged 12 to 16, attending secondary schools in economically deprived areas. Although BAM is a targeted selective intervention for students with social-emotional needs, it employs a strategic mixed-cohort recruitment model to avoid stigmatisation and the development of negative labels and promoting group safety and support relationship building within cohorts. Counsellors work closely with school leadership to recruit a balanced cohort comprising approximately 15% 'Tier 1' students (those who are thriving in their overall social and emotional development), 70% 'Tier 2' students (those facing some social and emotional challenges), and 15% 'Tier 3' students (those approaching or in crisis). This deliberate mix ensures a group dynamic where positive peer modelling can occur alongside targeted support. For the efficacy trial the cohorts will be ‘tiered’ using ‘risk to youth violence’ as a framework, instead of HSA tiers, a tiered system based on the Multi-Tiered System of Support. This is to ensure that YP with a high enough level of need are targeted. Furthermore, only boys in years 8-10 (age 12-15) will be recruited.

Delivery relies on full-time BAM Counsellors embedded within the school staff, who typically hold QCF-6 level qualifications in therapeutic or youth work fields. Crucially, these practitioners are recruited to be culturally relatable to the scholars, often sharing similar lived experiences or backgrounds, which is theorised to be essential for building the trust required for therapeutic work. To maintain fidelity to the model, counsellors undergo approximately 300 hours of programme training and receive robust "backbone support," including monthly in-person training, weekly peer supervision, and regular coaching from a "replication specialist" who assesses competencies such as clinical skills and youth engagement.

[21] Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S., and Pollack, H. A. (2017). "Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago." Quarterly Journal of Economics, 132(1), 1-54.

Programme content and adaptation

The intervention was originally designed to be delivered over two academic years through a multi-modal approach centering on "BAM Circles"—weekly, hour-long group sessions held during the school day with 8 to 12 scholars. The 2-year programme follows a 30-lesson licensed curriculum, the "BAMual," which utilises check-ins, role-playing, stories, and group missions to practice social-emotional skills. To reinforce this group work, counsellors conduct monthly "Brief Encounters"—informal check-ins lasting up to 15 minutes in communal school spaces—and provide One-to-One support for scholars with higher levels of need. Occasional "Special Activities," such as after-school sports or field trips, are also used to provide teachable moments in real-world environments. The target dosage for full implementation is 45 BAM Circle sessions over the two years, with attendance at a minimum of 13 sessions per year often defined as the threshold for sufficient exposure for evaluation purposes. BAM Circles facilitate youth reflection during regulatory changes, with Counsellors providing empathetic support to manage distress and address individual needs. The Counsellors facilitate a responsive environment that encourages self-reflection, accountability, and reintegration into the group.

This evaluation will focus on the evaluation of the new 1-year model adapted version. While retaining commitment to the six core values, the content of the lessons will be adapted to focus on the key drivers of internalisation of the values. As such, the two-year curriculum will be adapted to a one-year version which will retain many of the key activities. In keeping with the 2- year model, attending 13 BAM circle sessions in a given year is considered a minimum level of exposure to BAM for positive results to be observed, rather than being defined by a percentage of total lessons completed, or additional activities attended.

The adaptation of the programme to a 1-year model is intended to accommodate a funding landscape for SEL programmes in the UK where a single-year intervention may be more appropriate. We expect the willingness of philanthropic partners to make a commitment to supporting BAM to increase with a 1-year programme. Additionally, the BAM delivery team, having engaged dozens of potential local authority and school placement partners over the last six years, found that due to the budget setting cycles at these institutions, they signalled that one-year delivery commitments would be all they would be prepared to make when agreeing to resource and host the programme. As this did not align with the two year nature of the programme, many prospective delivery partners had to be turned away. This adaptation is expected to improve rates of adoption in England and Wales and ensure that the evidence base reflects the efficacy of the programme within these parameters. More details on the proposed approach to adaptation, including the principles that will be applied, and the timelines for this process, can be found in appendix 1. Further details about the one-year version of the BAM programme can be found in the TIDieR framework below (table 1).

We acknowledge that the adaptation to a one-year programme is a substantial shift and will ensure that this is thoroughly explored in the pilot year of the evaluation.

Table 1. TIDieR framework of the one-year BAM programme

TIDieR field	Guidance	Content
Why	Theory and rationale	BAM aims to improve school engagement and reduce involvement in youth violence through social-emotional learning. Its theory of change rests on a set of mechanisms: young people learn to self-regulate, build a relationship with a trusted adult, practise perspective-taking, gain a sense of achievement and belonging, and develop a pause before reacting to anger. These mechanisms are expected to improve self-esteem, self-control, and emotional regulation, which in turn should reduce exclusions, behavioural difficulties, and offending.
Who	Target population	Adolescent boys aged 12–15 (Years 8 to 10) living in areas of high need. Young people are selected if they experience challenges in resiliencies, relationships, or school engagement. Recruitment targets a mix of needs: 15% thriving, 70% balanced, and 15% in or approaching crisis.
What	Intervention activities	The programme consists of six activity types: **1. BAM circles**: weekly group sessions following a structured format (check-in,

		values-based activity, reflection, application). **2. After-school activities**: monthly sessions that consolidate the core value of the month. **3. Special activities:** group outings and celebrations outside of school time (e.g. affirmation trips, university visits, end-of-year celebration, intra-BAM football league)[22]. **4. Brief encounters:** informal, ad hoc check-ins (<15 mins), wherever and whenever counsellors see participants. **5. 1-2-1 support:** individualised, formally scheduled 30-minute counselling sessions for higher-level needs[23]. **6. Other activities**: proactive support around incidents, including visiting students after suspension/exclusion and accompanying them to the BAM room following an incident.
What	Materials	The primary material is the BAMual, a 25-lesson curriculum. Physical tools include boxing mitts, balls for group missions, video education (e.g., 'Jordan to the Max'), and emoji cards for reflection.
Who by	Delivery staff	Relatable male BAM

[22] Note that since these are not part of the core curriculum, they will not count towards dosage or fidelity. We will conduct exploratory analysis based on the frequency and intensity of these activities.

[23] Note that since these are not part of the core curriculum, they will not count towards dosage or fidelity. We will conduct exploratory analysis based on the frequency and intensity of these activities.

		counsellors with QCF-6 level qualifications. They are supported by a "backbone team" comprising Mental Health Foundation (MHF) and Youth Guidance (YG) staff.
How	Mode of delivery	Groups of 8–12 young people meet weekly for the BAM Circle. Delivery emphasizes "action and reflection," using experiential learning, archetypes, and goal setting.
Where	Setting	Mainstream secondary schools and initially a Pupil Referral Unit (PRU). Sessions occur in a dedicated school room or communal spaces like hallways/playgrounds for brief encounters.
When and how much	Dosage	The programme lasts one school year. It includes 27 one-hour sessions during school and 6 after-school sessions after each core value. Brief encounters are targeted at once per month per scholar.
Tailoring and flexibility		Surface adaptations for the UK context included changing "city blocks" to "streets" and "basketball" to "football". Deep adaptations replaced "tribal" with "ancestry" and changed the archetypes of "savage/warrior" to "destructive/constructive" to avoid problematic connotations.

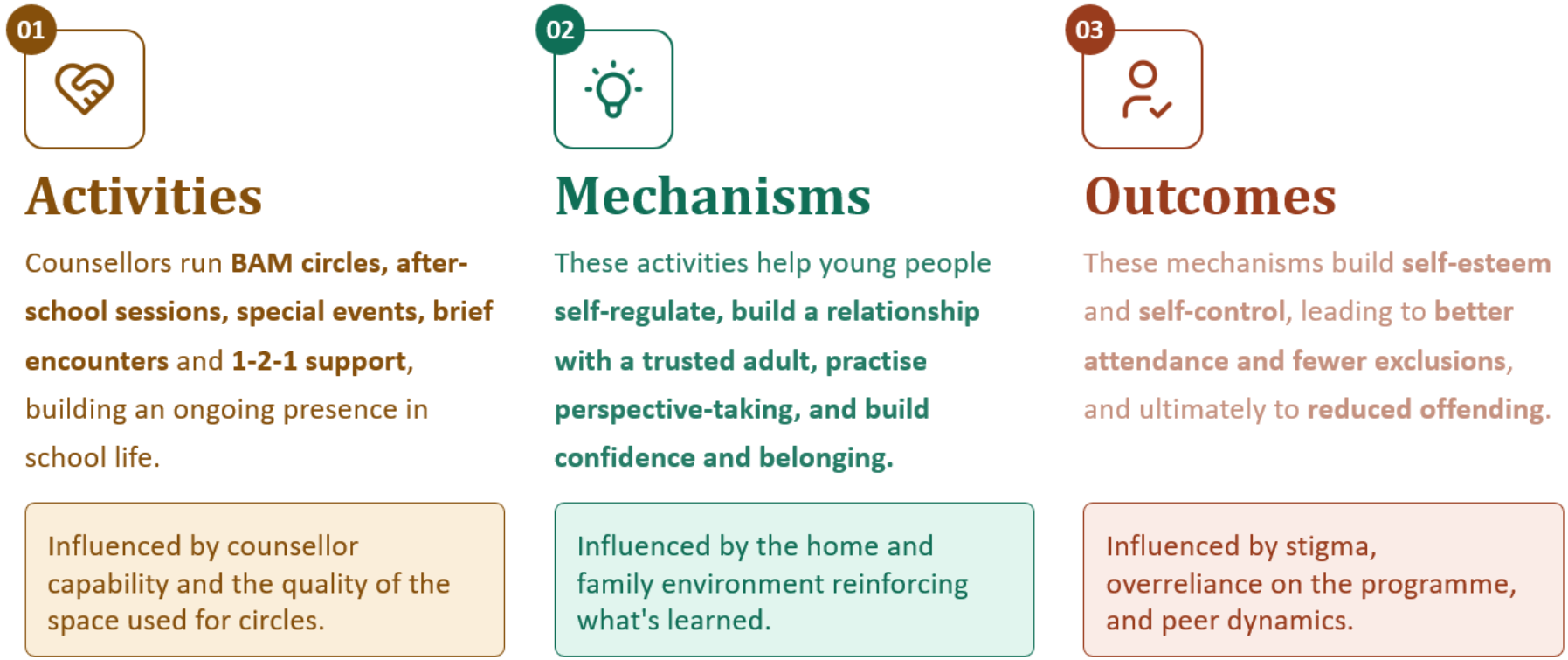
Monitoring and fidelity		Monitoring and fidelity is tracked via counsellor implementation checklists, 25-lesson tracking, and a 6-step coaching model that assesses five core counsellor competencies on a 5-point metric.

BAM Theory of Change

The intervention is hypothesised to work through a mechanism of "action and reflection" facilitated within a safe, consistent therapeutic relationship. The programme theory posits that by internalising six core values—including Integrity, Accountability, and Self-Determination— boys develop "positive assets". These assets serve a dual function: buffering young people from risks in their environment and empowering them to capitalise on opportunities and resources, such as education, thereby facilitating responsible decision-making and a positive transition to adulthood. It aims to build social-emotional capacity—specifically emotional regulation and the ability to read social cues—by fostering supportive relationships that encourage open and honest communication. This relational foundation provides a safe space for boys to express and understand their emotions, acting as a protective factor against the frustration and emotional isolation that can disrupt regulation, ultimately leading to short-term reductions in school-based behavioural incidents. Over the medium to long term, this improved regulation and internalisation of Self-Determination is expected to lead to more responsible decision-making and translate into higher academic attainment and increased graduation rates. These, along with the value of Visionary Goal Setting would lead to a reduction in maladaptive behaviours, including a reduction in violent crime arrests.

Figure 1 below presents the high-level theory of change for the one-year programme. The full theory of change can be found in appendix 2.

Figure 1: High-level theory of change



Research questions or study objectives

Pilot

The overarching question the **pilot trial** will assess is:

> **Is there sufficient evidence on the evaluability, deliverability and acceptability of Becoming A Man, for the project to progress to a full-scale efficacy trial?**

The pilot is also designed to be an **internal** pilot (Avery et al., 2017). This means that as part of the pilot trial, outcomes will also be measured (i.e. collect high-quality outcome data that could be analysed alongside outcome data collected in any subsequent efficacy trial phase to determine intervention impact).

It is primarily designed to achieve the following specific objectives:

- **Objective 1: To assess whether there is sufficient evidence that a full-scale efficacy trial is feasible in this context** (particularly in terms of recruiting sufficient numbers of higher vulnerability young people into the evaluation, randomisation, avoiding spillovers between the treatment and control arms and outcome data collection) to justify extending the sample and continuing to efficacy trial.
- **Objective 2:** To collect high-quality outcome data as part of the pilot trial that could be analysed alongside outcome data collected in any subsequent efficacy trial. The sample size in the pilot alone is unlikely to be sufficiently large to make strong causal claims about effectiveness.
- **Objective 3:** To monitor programme deliverability, including assessing participant recruitment, randomisation and retention rates. This is to evaluate delivery fidelity and quality under the 1 year model and identify implementation barriers and enablers.
- **Objective 4:** To assess the acceptability of the BAM programme and trial arrangements (e.g. randomisation, data collection process and use of local police data) among participants, school staff and practitioners across diverse cultural and ethnic backgrounds.
- **Objective 5:** To empirically estimate and confirm the pre-post test correlation on the primary outcome measure (SDQ total) to determine whether refinements to statistical power assumptions or sample size calculations are required prior to the full efficacy trial.

This learning will inform any necessary refinements made between the pilot and efficacy stages, which may include refining referral pathways and recruitment processes, refining

how the project (programme and evaluation) is communicated to young people and schools, how the programme is delivered, and key evaluation procedures.

Efficacy Trial

The overarching question that the **efficacy trial** will address is:

> **Among boys in year 8-10 with high-risk secondary or tertiary levels of need, does a 1 year school-based social and emotional learning programme designed to help children engage in responsible decision-making, reduce behavioural difficulties, compared with business as usual**

The primary research question is focused on boys who are assessed as having high-risk secondary or tertiary levels of need. In order to maintain fidelity to the model, the programme must be delivered to a mixed cohort, including around 15% of the group being of low risk. Limiting the primary analysis to those assessed as higher risk ensures that fidelity is maintained while focusing on the most relevant population, increasing precision in estimates and narrowing confidence intervals.

It is primarily designed to**:**

- **Establish impact** - Among boys eligible for BAM, does a one-year school-based social-emotional learning programme designed to build self-regulation and trusted-adult relationships reduce behavioural difficulties at post-intervention, relative to business as usual?
- **Establish mechanisms and moderating factors** - is there evidence of differential impact according to sample characteristics, such as home environment or existing relationships with adults at school? Where statistically significant differences between groups are identified, does the pattern of results conform with the expected sequence from short-term gains (self-esteem, self-control, emotional regulation) to medium-term gains (attendance, reduced exclusions) to the long-term outcome of reduced offending?
- **Further understand key aspects of implementation -** including reach, quality and fidelity of delivery, and acceptability. The main adaptation under test is duration: BAM is usually delivered over two school years, and this trial delivers a compressed one-year version. What does that compression mean for dosage and fidelity, and what curriculum or activity adaptations are made as a result? Are BAM and the trial arrangements acceptable to young people, and does acceptability vary across young people from different cultural and ethnic backgrounds?

Please see the full set of research questions below in Table 2.

Table 2: Specific research questions

Project phase	Research objective	Focus	Research questions
Pilot Trial	Establishing **evaluability**	RQ 1: Is there sufficient evidence that study recruitment can reach the need of the efficacy trial?	*Referrals* RQ 1.1: How many young people are referred to the project/study over the pilot trial period? RQ 1.2: What factors affect the volume of referrals? *Referral suitability* RQ 1.3: What proportion of referred young people are eligible for the programme? RQ 1.4: What are the most common reasons for young people being deemed ineligible for the project/study? RQ 1.5: Do the referral criteria create a cohort with a sufficient proportion of highly vulnerable boys, as defined as 85% of participants being assessed as being either upper secondary or tertiary level of need? How do the HSA tiers map onto the YEF levels of need in our sample? *Recruitment and consent*

			RQ 1.6: How many eligible young people can be successfully consented into the evaluation? RQ 1.7: How does this vary across participant characteristics, including: ○ Ethnicity ○ School year ○ Free school meal eligibility RQ 1.8: Given the number of eligible young people successfully consented into the evaluation over the time-period of the pilot, would a well-powered RCT be achievable at the efficacy stage over the currently planned time period?
		RQ 2: Is there sufficient evidence that the randomisation process reaches the need of the efficacy trial?	*Feasibility and adherence* RQ 2.1: Is the randomisation approach feasible (i.e. is the randomisation straightforward as planned in this context, or are there unexpected barriers)? RQ 2.2: How many/what proportion of recruited

			participants complete baseline data surveys and are randomised? RQ 2.3: Is randomisation adhered to (i.e. is randomisation accidentally or intentionally subverted)? RQ 2.4: How could the approach to randomisation be adapted to increase feasibility and adherence? *Acceptability* RQ 2.5: Is randomisation acceptable to young people? RQ 2.6: Do BAM practitioners feel that evaluation activities (randomisation) impact the ability to deliver the programme well? In what way? RQ 2.7: What factors affect acceptability of randomisation? RQ 2.8: How could the approach to randomisation be adapted to increase acceptability?

		RQ 3: Is there any evidence that control group services are sufficiently similar to BAM in a way that may pose problems for the efficacy trial?	RQ 3.1: What alternative services or support do the control group receive over the trial period, if any? RQ 3.2: To what extent are these similar to Becoming a Man?
		RQ 4: Are the proposed data collection procedures successful and suitable for use in the efficacy trial?	*Participant perceptions of data collection and, where data allows, how does this vary by key characteristics* RQ 4.1: How do participants feel about the questions asked in the outcome data survey? RQ 4.2: How do participants feel about the length of the outcome data survey? RQ 4.3: How do participants feel about having Local Police Data used in the evaluation? RQ 4.4: How do participants feel about the questions asked and approach taken in depth interviews? *Attrition rates and retention in evaluation* RQ 4.5: What proportion of responses, particularly the

			primary outcome, are missing at each time point? RQ 4.6: How many/what proportion of treatment group participants complete post-programme surveys? RQ 4.7: How many/what proportion of control group participants complete post-programme surveys? RQ 4.8: What data collection approaches work well in retaining treatment and control group participants and, where data allows, do they differ by participant characteristics? RQ 4.9: What factors affect attrition rates? Do survey attrition rates vary by participant characteristics? ○ Ethnicity ○ Age ○ Free school meal eligibility ○ Baseline scores on outcome measures ○ Young persons experience of

			discrimination as measured by the Everyday Discrimination Scale
		RQ 5: Is there sufficient evidence that effect sizes and sample size meet the need of the efficacy trial?	RQ 5.1: For each outcome, what is the point estimate of effect size, what is the confidence interval around it, and what implication would the range of plausible values have for the required sample size at the efficacy stage? RQ 5.2: Is there any evidence of i) significant interaction between participants in the treatment and control arm ii) significant interaction between practitioners and control group participants or iii) spillovers from the whole-school approach which creates risk of contamination and will significantly impact our ability to detect an effect?
			RQ 6: Is it possible to collect the data that would be required to assess whether outcomes vary by

			key characteristics? ○ Ethnicity ○ Age ○ Baseline outcomes ○ Free school meal eligibility ○ Attendance/engagement with the programme ○ Young persons experience of discrimination as measured by the Everyday Discrimination Scale
	Measuring **outcomes**		RQ7.1 (impact): What is the impact of the BAM programme, compared to BAU, on pre-specified outcomes for boys in the study at the post-programme timepoint and 4-month follow-up? RQ7.2 (estimation): For each

			outcome, what are the estimated effect sizes, confidence intervals, and directions of effect? **Primary analysis:** What is the impact of BAM, compared to BAU, on behavioural difficulties among boys assessed as upper secondary or tertiary level of need at the post-programme timepoint and the 4-month follow-up? **Secondary analyses:** What is the impact of BAM, compared to BAU, on the following outcomes? Among boys assessed as upper secondary or tertiary level of need: o Arrests Among all boys: o Emotional and behavioural difficulties o Self-esteem

			○ Self-control and emotion regulation ○ Social skills ○ School attendance ○ Exclusions and suspensions ○ Externalising score ○ Arrests (exploratory)
	Monitoring **deliverability**	RQ 8: Is there sufficient evidence that programme recruitment and take up reach the need of the efficacy trial?	*School Recruitment* RQ 8.1: Can the delivery team recruit a sufficient number of schools to participate in the programme? RQ 8.2: What is the conversion rate for schools? I.e. How many schools who are approached about the programme take it up? ○ What are reasons schools provide for not wanting to take part? *Young people* RQ 8.3: How many eligible young people randomised to receive

			BAM initially take it up[24]? RQ 8.4: Does take-up vary by: ● Ethnicity ● Age ● Free school meal eligibility ● Baseline scores on outcomes. ● Young persons experience of discrimination as measured by the Everyday Discrimination Scale RQ 8.5: Can the delivery partner (via monitoring of school's recruitment efforts) recruit a sufficient proportion of CYP from minority ethnic communities (40%) into the project/study?
		RQ 9: Is there sufficient evidence that programme completion can reach the need of the efficacy trial?	RQ 9.1: How many CYP randomised to BAM attend each session RQ 9.2: How many complete BAM: defined as attending at

[24] Take-up is defined as attending at least one programme session. Dosage and compliance are defined differently, see below for more details.

			least 13 BAM group sessions? RQ 9.3: Does this vary by ethnicity, age, and baseline scores on key outcomes?
		RQ 10: Is there sufficient evidence that the level of fidelity the programme is being with meets the needs of the efficacy trial?	RQ 10.1: What adaptations to the 2-year model have been implemented? What has been kept, what has changed? ○ What are the views on the adaptations? RQ 10.2: Is the programme being delivered with fidelity to the 1-year model? If not, why not? RQ 10.3: What are the barriers and facilitators to delivering the programme well and with fidelity? RQ 10.4: What variations in delivery are appropriate for effective implementation? RQ 10.5: What variations in delivery are not appropriate for effective implementation? RQ 10.6: Does the level of fidelity differ across counsellors and schools? RQ 10.7: Is the training process used in the evaluation consistent

			with standard practice?
			RQ 11: What is the average cost of delivering BAM?
	Monitoring **acceptability**[25]	RQ 12: Is there enough evidence that programme acceptability can reach the need of the efficacy trial?	RQ 12.1: Is BAM acceptable to young people and schools? If not, why not? RQ 12.2: Are there any barriers/facilitators of programme acceptability? RQ 12.3: Do views on acceptability of the intervention differ amongst different groups of young people, with a specific focus on young people from racially minoritised backgrounds where sample size allows. RQ 12.4: Does acceptability vary depending on how the programme is delivered?
Efficacy Trial	Establishing Impact		RQ13.1 (impact): What is the impact of the BAM programme, compared to BAU, on pre-

[25] Acceptability refers to the extent to which participants and practitioners judge the intervention to be appropriate, engaging, and valuable. It encompasses both their experience of taking part (e.g., enjoyment, comfort) and their perceptions of its usefulness (e.g., whether the programme feels helpful, relevant, and likely to have a positive impact).

			specified outcomes for boys in the study at the post-programme timepoint and the 4 month follow-up? RQ13.2 (estimation): For each outcome, what are the estimated effect sizes, confidence intervals, and directions of effect? **Primary analysis:** What is the impact of BAM, compared to BAU, on behavioural difficulties among boys assessed as upper secondary or tertiary level of need at the post-programme timepoint and 4 month follow-up? **Secondary analyses:** What is the impact of BAM, compared to BAU, on the following outcomes? Among boys assessed as upper secondary or tertiary level of need:

			○ Arrests Among all boys: ○ Emotional and behavioural difficulties ○ Self-esteem ○ Self-control and emotion regulation ○ Social skills ○ School attendance ○ Exclusions and suspensions ○ Externalising score ○ Arrests (exploratory)
	Establishing mechanisms and moderating factors		RQ 14: Is there evidence of differential impact according to specific sample characteristics? Where sample size allows, does descriptive analysis support evidence of intersectionality across groups? The sample characteristics that will be explored individually and interactionally include: 1. Ethnicity 2. Age

			3. Free school meal eligibility 4. Baseline scores on outcomes (including HSA) 5. Short-form "Everyday discrimination scale" 6. If statistically significant differences between groups are identified, do the pattern of results conform with expected short-, medium- and long-term outcomes?
	Assessing acceptability	RQ 15: Is the programme acceptable to stakeholders?	RQ 15.1: Is BAM acceptable to young people and schools? If not, why not? RQ 15.2: Are there any barriers/facilitators of

			programme acceptability? Are these the same/similar to those expressed during the pilot trial? RQ 15.3: Do views on acceptability of the intervention differ amongst different groups of young people, with a specific focus on young people from racially minoritised backgrounds. RQ 15.4: Does acceptability vary depending on how the programme is delivered? RQ 15.5: What is the perceived impact of the programme? RQ 15.6: What are the perceived mechanisms creating the programme's impact?
	Assessing deliverability	RQ: 16: Is the programme being delivered with fidelity and what factors impact fidelity?	RQ 16.1: Is the programme being delivered with fidelity? If not, why not? RQ 16.2: What are the barriers and facilitators to delivering the programme well and with fidelity? RQ 16.3: What variations in

			delivery are appropriate for effective implementation? RQ 16.4: What variations in delivery are not appropriate for effective implementation? RQ 16.5: Does the level of fidelity differ across different practitioners? RQ 16.6: Does the level of fidelity differ for those schools that have been included in the pilot and efficacy trial compared to those that are only involved in the efficacy phase? RQ 16.7: Does the level of fidelity vary in the schools who have already delivered BAM?

Progression criteria

Monitoring and progression criteria will be used throughout the project for two purposes:

- To monitor if the project is proceeding as expected, allowing for us to make adjustments or pause the work if needed.
- To make recommendations to the Youth Endowment Fund as to whether progression to an efficacy trial should be pursued at the end of the pilot study. Recommendations on progression will be made in two stages. The first will take place before the end of the first school year, in April. This will cover all progression criteria for which data exists at this point. The second stage will take place in the summer, in August, and will cover all remaining criteria. The table below indicates which stage each criteria is covered in.

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

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

The quantitative monitoring criteria used to monitor evaluability-related objectives are described in Table 3 below. While the criteria below offer guidance for the progression of the evaluation on the basis of quantitative assessments, these will also be complemented by qualitative measures, such as ongoing practitioner feedback and interviews with practitioners and caregivers. Overall, a deterministic rule about how many green-, amber-, and red-rated criteria would justify/prohibit progression will not be established, but instead will use the criteria to support a balanced judgement, weighing the importance of each criteria in the round.

Table 3: Progression criteria

Research objective	Criterion	Description	Target	Measurement	RAG scores	Review point
Deliverability	School recruitment (pilot)	The number of schools that are recruited, consented and retained in the pilot trial	3	Delivery partner data	Green: 3 Amber: 2 Red: 0-1	Stage 1: April
	School recruitment (efficacy)	The number of schools that have agreed to participate in the efficacy trial as evidenced by an agreement in principle	6	Delivery partner data	Green: 6 Amber: 4-5 Red: 3 or fewer	Stage 1: April
	Eligible referral volume	The number of eligible referrals received across all schools as a proportion of total intended sample	~ 90 per delivery year per school	School referral forms	Green: 75-100% Amber: 60-74% Red: less than 60%	Stage 1: April
	Higher-risk CYP	The proportion of	-	School referral	Green: 85-100%	Stage 1: April

	enrolment	CYP who enter the trial who are classified as secondary level or tertiary level of need		forms	Amber: 70-84% Red: less than 70%	
	Enrolment in the programme	Proportion of referrals who agree to participate in the project, and provide baseline data and consent to randomisation	-	EYV Lab database	Green: 75-100% Amber: 60-74% of Red: less than 60%	Stage 1: April
	Diversity of sample	Proportion of young people consenting to participate in the evaluation who are from minority ethnic backgrounds	-	School referral forms	Green: 40-100% Amber: 20-39% Red: less than 20%	Stage 1: April
	Dosage	Proportion of young people who receive minimum	-	Practitioner/supervisor reported data	Green: 75-100% Amber: 50-74% Red: less than 50%	Stage 2: August

		level of input[26]				
Acceptability	Acceptability to young people	Proportion of young people who report being happy with the programme, i.e. they are satisfied or very satisfied with the programme	-	Endline young person survey	Green: 75-100% Amber: 50-74% Red: less than 50%	Stage 2: August
	Acceptability to schools	Proportion of school staff who report being satisfied or very satisfied with the programme	-	Endline school staff survey	Green: 3 schools Amber: 1-2 schools Red: 0 schools	Stage 2: August
	Acceptability to practitioners	Proportion of practitioners who report being satisfied or very satisfied with the	-	Endline practitioner survey	Green: 3 practitioners Amber: 1-2 practitioners Red: 0	Stage 2: August

[26] Defined as attending 13 sessions, the minimum dosage threshold for starting to show positive outcomes, per the US studies.

		programme			practitioners	
Evaluability	Completeness of baseline survey data collection	The proportion of high-risk respondents[27] answering at least 80% of the questions.	-	Baseline young person survey	Green: 85-100% Amber: 70-84% Red: less than 70%	Stage 1: April
	Evaluation retention - treatment	Proportion of high-risk participants randomised to the treatment group who complete the post-intervention survey.	-	Post-programme young person survey	Green: 90-100% of participants randomised to treatment Amber: 70-89% Red: less than 70%	Stage 2: August
	Completeness of post-intervention data collection - treatment	The proportion of high-risk respondents answering at least 80% of the questions	-	Post-programme young person survey	Green: 85-100% Amber: 70-84% Red: less than 70%	Stage 2: August

[27] Defined as secondary level or tertiary level of need

	Evaluation retention - control	Proportion of high-risk participants randomised to the control group who complete the post-intervention survey.	-	Post-programme young person survey	Green: 90-100% of participants randomised to control Amber: 70-89% Red: less than 70%	Stage 2: August
	Completeness of post-intervention data collection - control	The proportion of high-risk respondents answering at least 80% of the questions.	-	Post-programme young person survey	Green: 85-100% Amber: 70-84% Red: less than 70%	Stage 2: August

Impact evaluation: Design

This pilot is designed as a two-armed individually randomised RCT, with the treatment arm receiving Becoming a Man, and the control group receiving business as usual services. The BAM programme will be delivered between October and July of the delivery school year.

The delivery period evaluated by the pilot trial is designed to last one school year. The delivery period evaluated by the efficacy trial would last for two school years, and follow a similar design and process.

To ensure that there are a sufficient number of eligible boys each year in each school, school's involvement will be limited to two years. As such, delivery will take place as follows. Schools will run 3 circles per year with 12 pupils randomised to each circle. To ensure that the schools have sufficient numbers of eligible pupils, schools who have recently completed a cycle of BAM for at least the first year will be excluded. This is to give them sufficient time to rebuild an eligible pool, without having a younger or less high risk cohort, as the current eligible population will already have been offered BAM.

While the current plan is to deliver the programme across nine schools, the delivery team would have the budget to recruit an additional counsellor, allowing a tenth school to be included in the study during the efficacy phase, in years two and three. This will only take place if it is necessary to achieve the sample size and will be reviewed during the pilot year, depending on the total number of referrals that progress to randomisation. If recruitment in the pilot year falls below 80% of the target (i.e. 86 boys), we anticipate that the 10th school will be required to achieve the necessary sample size.

	Year 1 (pilot)	Year 2 (efficacy)	Year 3 (efficacy)
School 1			
School 2			
School 3			
School 4			
School 5			
School 6			
School 7			
School 8			
School 9			

(Potential) School 10			
Total number of circles	**9**	**18**	**18**
Total number of boys	**108**	**216**	**216**

The planned methodologies for the pilot and efficacy phases are largely identical, but simply repeated on a larger-scale for the efficacy phase.

Table 4: Trial design

Trial design, including number of arms		Two-arm individual-level RCT (internal pilot followed by extended efficacy phase)
Unit of randomisation		Individual young people
Stratification variables (if applicable)		HSA tiers - to ensure balance in each group
Primary outcome	variable	Behavioural difficulties only for those assessed as being at upper-secondary or tertiary level of need
	measure (instrument, scale, source)	Young person self-report on the Strengths and Difficulties Questionnaire total score, measured post-intervention.
Secondary outcome(s)	variable(s)	Arrests only for those assessed as being at upper-secondary or tertiary level of need

	measure(s) (instrument, scale, source)	A binary variable indicating whether a boy has been arrested for any offence (violent and non-violent) in the 18 months following randomisation, using local police data.
	variable(s)	Behavioural difficulties for all students
	measure(s) (instrument, scale, source)	Young person self-report on the Strengths and Difficulties Questionnaire total score, measured post-intervention
	variable(s)	Pro social behaviour
	measure(s) (instrument, scale, source)	Young person self-report on the prosocial scale of the Strengths and Difficulties Questionnaire, measured post-intervention
	variable(s)	Externalising behaviour
	measure(s) (instrument, scale, source)	Young person self-report on the externalising scale of the Strengths and Difficulties Questionnaire, measured in post-intervention
	variable(s)	Self-esteem

	measure(s) (instrument, scale, source)	Young-person self-report on the Rosenberg Self-esteem Scale, measured post-intervention.
	variable(s)	Self-control and emotional regulation
	measure(s) (instrument, scale, source)	Young person self-report on the Difficulties in Emotion Regulation Scale, measured post-intervention.
	variable(s)	School attendance
	measure(s) (instrument, scale, source)	Percentage attendance in the year following the programme
	variable(s)	Exclusions, suspensions and transfers
	measure(s) (instrument, scale, source)	A binary measure capturing the likelihood that the CYP faced either exclusion, suspensions or transfers (faced any of the three issues = 1; 0 otherwise). Combined measure as likelihood of any of these events is low.
Baseline for primary and secondary outcomes	The baseline for outcomes measured using surveys is the same measure, collected prior to randomisation. The baseline for non-survey measured outcomes is the same measure for the same period prior to randomisation (i.e. for arrests it is 18 months prior to randomisation and for school data it is over the school year prior to	

	randomisation. .

Randomisation

The pilot trial will use a within-schools randomisation design, with randomisation taking place at the level of the individual student. Stratified randomisation will be used to ensure that participants within the treatment group and control group are balanced on HSA. The HSA evaluates a pupil's entire development by considering cognitive, social, emotional and ethical growth. As part of our pilot trial it will be established how the HSA maps on YEF's tiers of vulnerability. Pupils will be classified into the HSA tiers based on their baseline survey. This is to ensure there is a sufficient balance within the treatment group to facilitate mixed-tier intervention groups, which is a core component of the programme model as well as ensuring balance between treatment and control.

Randomisation at the individual level is being proposed, in line with all previous US evaluations of BAM involving randomisation, since the estimated risk of contamination is low. While there is a school-wide component of BAM, this is only considered to be a facilitator to successful delivery of BAM. The key driver of change are the BAM circles, which only treated individuals will receive. Additionally, since delivery will only be in schools for two years, and the school wide- component takes time to build, sufficiently large levels of this to lead to contamination are not expected. This will also be monitored as part of the fidelity measure, and so the extent of contamination will be captured and reviewed during the pilot. The evaluators are working with the delivery team to maintain a live document that rates various activities counselors might do in schools in terms of their risk to contamination. See Appendix 5 for the current version.

Randomisation and assignment to intervention or control groups will be implemented by the Ending Youth Violence Lab using a random number generator in Stata. The allocation mechanism will be young people's Unique Pupil Number (UPN). Randomisation will occur once for each school after students are referred to the programme and assessed for eligibility, and after both parental and child consent has been obtained and baseline assessment has taken place, but before any structured contact with the programme. To ensure transparency, the randomisation process will be fully documented using timestamped Stata log files, and the syntax will be quality assured by a senior EYVL colleague prior to execution. It will not be possible to blind delivery staff, EYVL analysts conducting the impact evaluation, or the young people themselves to allocation. The master link key connecting UPNs to treatment status will be securely stored in a restricted, password-protected folder. During analysis, researchers will only have access to a masked dataset (e.g., coded as Group A and Group B) and will not have access to the link key. Unblinding will only occur once the primary statistical analysis is

complete, or in the event of a critical safeguarding incident requiring knowledge of group allocation.

This means that after the delivery team has received referrals, assessed eligibility and parental consent has been received, the Ending Youth Violence Lab will conduct baseline data collection and randomise consented students to either treatment or control. Randomisation will be conducted at the individual level within each school. To ensure a balanced allocation and facilitate the formation of mixed-tier BAM circles, randomisation will be stratified by school and HSA tier. The Ending Youth Violence Lab will then inform the delivery team about the allocation of young people to treatment or control. Individuals allocated to the control group will receive BAU. In previous BAM delivery schools, BAU varied by borough and by school site, with a majority using the PSHE curriculum to meet statutory requirements addressing safety, health, and inclusion and these were often delivered through a combination of discrete lessons, assemblies, and external partnerships with organisations like Football Beyond Borders, Kooth, and/or Place2Be. Though the PSHE offer varies by school site, the general framework focuses on equipping young people with resilience and modern life skills - all of which complement BAM. For schools involved in the RCT, the delivery team is working on compiling the BAU offer available at each school site and these details will be included in the final protocol. As part of data collection, it will be investigated what this entails through surveys of control group participants.

Participants

Process

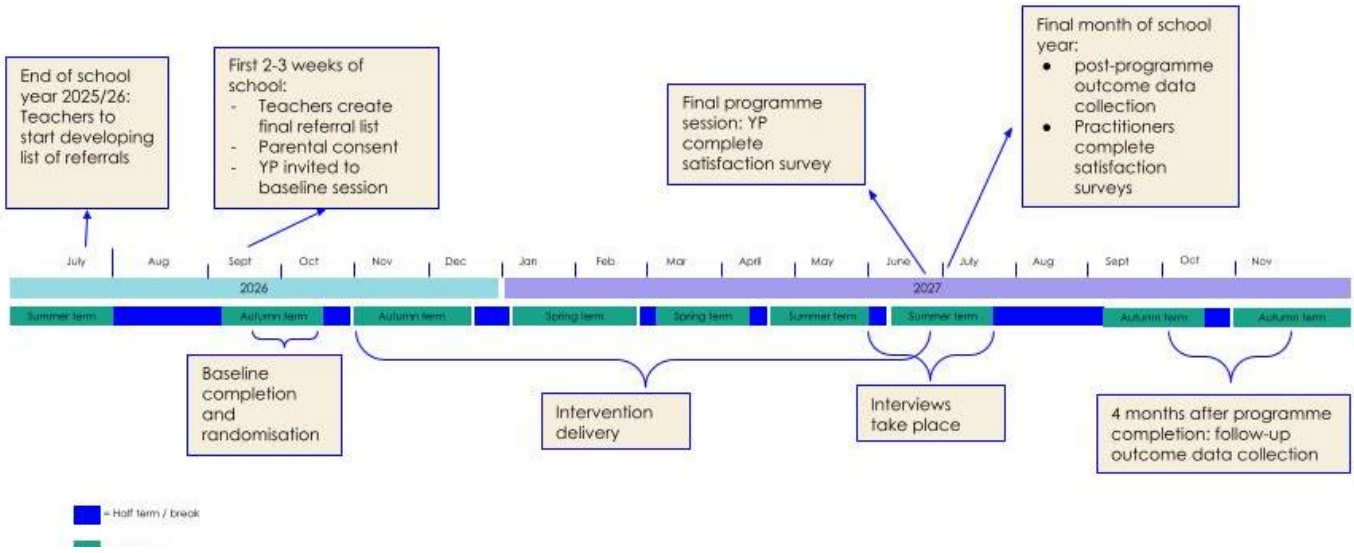


Figure 2. Research process steps and timeline (blue: half-term/break, blue: term time).

Our evaluation approach for the BAM programme follows a structured timeline integrated within the school calendar. The process maps out a clear journey for participating children, school staff, and researchers across the following key steps:

- **Promotion and school engagement**: BAM counsellors initiate the process by hosting targeted workshops for relevant school leaders, heads of year, and pastoral teachers in years 8–10. These sessions are designed to inform school staff about the evaluation design, the specific multi-tiered eligibility criteria, and the structural requirements for generating student referrals.
- **Referral generation**: The identification of eligible boys relies on close cooperation between the school pastoral team, the special educational needs coordinator (SENCO), and the embedded BAM counsellor. Teachers are to begin thinking about referrals at the end of the preceding school year in July.
- **Eligibility assessment**: Upon receiving the referral lists, the delivery team reviews the referred young men against the defined inclusion and exclusion protocols. To be deemed eligible, participants must be boys in academic years 8–10 who present with a demonstrable lower secondary, upper secondary, or tertiary level of social-emotional need or risk related to youth violence.
- **Parental consent and onboarding**: During the initial weeks of the autumn term, the school and delivery partners secure formal parental consent for the eligible students. This will cover consent for both the evaluation and the programme (using two different forms). Once parental consent is obtained and logged, the boys are formally onboarded into the evaluation process and invited to participate in the baseline data session. This will include consent for both the collection of police data and data archiving.
- **Baseline data collection**: Prior to randomisation, the Ending Youth Violence Lab conducts baseline data collection. Surveys are administered in person within a designated school room using tablets or computers, with an evaluation research team member present to provide support tailored to the students' literacy and regulatory needs.
- **Randomisation**: Immediately after baseline assessments are complete and all consent forms are verified, the Lab executes individual-level randomisation. Using a random number generator in Stata linked to the student's Unique Pupil Number (UPN), boys within each school are allocated on a 1:1 ratio to either the treatment or control arm. This randomisation is stratified by HSA tiers to guarantee balanced group dynamics and clear baseline equivalence.
- **Allocation communication and delivery**: The Lab communicates the final allocations to the delivery team. BAM counsellors engage the treatment group to begin intervention delivery, which features weekly BAM Circles alongside ad-hoc brief encounters, one-to-one support, and special activities spread across the school year.

Boys allocated to the control group do not receive the intervention and instead receive standard BAU .

- **Post-programme data collection**: In the final month of the school year (June and July), the evaluation team administers the post-programme endline surveys to all participating young people in person. Treatment group participants complete a short satisfaction survey during their final BAM session, while counsellors complete satisfaction surveys at the end of the cohort. Simultaneously, qualitative research takes place via in-depth interviews with a sample of youth, practitioners, school staff, and caregivers to explore trial acceptability and feasibility.
- **Long-term outcome and administrative tracking**: Follow-up survey data is collected from the young people approximately 4 months after programme completion (Month 13 post-randomisation). To track long-term trajectories, trial data is linked with local police records to evaluate arrest and offending outcomes 18 months post-randomisation. Finally, school-level administrative records are collected at Month 21 post-randomisation to objectively evaluate long-term impacts on school attendance, suspensions, and exclusions.

Referral

BAM counsellors work directly with senior school leadership, heads of year, and school pastoral teams to identify eligible boys. Staff identify boys who are experiencing challenges in their social-emotional development and might benefit from the programme. Counsellors guide staff to refer a mixed cohort of students representing the three "tiers of need" (thriving, balanced challenges/strengths and approaching crisis).

The recruitment of boys will take place in September and October of each school year.

Inclusion and exclusion criteria for young people

Our aims in defining inclusion and exclusion criteria are to:

1) Include young people who stand to benefit most from participation in the project

2) Exclude any young people for whom the project and/or programme may not be best suited, or may be at an increased risk of harm if they were to participate.

<u>Inclusion criteria</u>

The vast majority of boys must meet one of the following criteria (i.e. must present with either a 'lower secondary', 'upper secondary' level, or a 'tertiary' level of need). Schools may refer a small number of boys who do not meet one of the following criteria if deemed appropriate in conjunction with the counsellors. This would be to ensure balance across the circles.

Once boys are referred, they cannot take part in subsequent years. Boys must be in years 8-10 to take part.

- Lower secondary level of need:
    - Are flagged by school staff (e.g., pastoral lead, form tutor) as needing general support to improve behaviour and/ or social-emotional skills, despite generally positive academic and behavioral records.
    - Have a history of minor, infrequent low-level disruptive or challenging behavior that did not result in a fixed-term exclusion.
    - Exhibit emerging emotional difficulties (e.g., low mood, anxiety, minor conflicts with peers).
    - Are experiencing a minor or temporary dip in academic engagement or attainment that staff attribute to a social or emotional factor.
    - Have multiple in-school disciplinary incidents in the past year indicating high risk of escalating risk.
- Upper secondary level of need:
    - Have received one or more fixed term exclusions
    - Have serious and frequent emotional or behavioural difficulties (e.g substantial conflicts with peers, elevated anxiety, depressive symptoms, persistent rule breaking)
    - Have been identified as vulnerable to involvement in gangs or are known to be gang affiliated
    - Are known to have substance abuse/misuse problems
    - Have been identified as at risk of exploitation, by adults or by criminal peers.
    - Have a sibling(s) or parent that has entered into the criminal justice system.
- Tertiary level of need:
    - Have been arrested and received a caution or warning from the police
    - Have received a diversion or out of court disposal
    - Are receiving services from a Youth Offending Team

Many of these categories are subjective, creating a risk that referrals are subject to unconscious biases and cultural stereotyping. However, using more objective measures such as attendance or achievement does not fully mitigate these issues and also means that cohorts may be created who are some distance from involvement in serious violence. To address these challenges the Lab will work with delivery partners, ClearView and YEF's REA to develop case studies and personas for schools illustrating the sorts of behaviours which would and would not make an individual eligible for BAM (see Appendix 3 for examples). These resources will also encourage staff to reflect on what behaviours may be viewed through a racialised lens, and would not be applicable here. We will also include a free text box in the referral form for referrers to explain a referral if the criteria are not appropriate.

Based on analysis conducted by Islington, they are confident that a substantial proportion of the eligible population presents with vulnerability factors consistent with BAM's target group and the evaluation's focus on youth violence prevention. The schools that are likely to take part have rates of attendance, persistent absences, suspensions and exclusions that are in line with delivering substantially high numbers of eligible students.

Group composition

While individual eligibility is defined above, the BAM Circle model relies on a balanced cohort to function effectively. To ensure the programme is delivered to a cohort with sufficient levels of risk every BAM Circle must be at least 8-9 pupils (85% across the entire sample) assessed as being at an 'Upper Secondary' or 'Tertiary' level of need. The distribution of level of need will be monitored as referrals come in to ensure we are meeting this requirement.

Exclusion criteria

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

- Are already participating in intensive counselling, therapy or socio-emotional training.
- Plan to move out of the local area within the delivery period, and thus may not be available for full delivery of the programme.
- Have severe developmental delay which may prevent participating in the intervention and evaluation, including inability to practice abstract thinking
- High level of chronic absenteeism in previous year.
- Are actively homicidal
- Are actively suicidal or psychotic.

Inclusion and exclusion criteria for schools

Inclusion criteria are:

- Mainstream schools for secondary school students

- Estimate that they would be able to refer at least 90 boys across years 8-10 who are eligible for BAM in each year of delivery (see Appendix 4 for the recruitment and retention table)
- Able to commit to supporting the evaluation activities

Exclusion criteria are:

- A school that offers a similar intervention
- For year one, schools who have recently finished delivering BAM will be excluded

Sample size calculations

The ideal number of boys assigned to each circle is 12. This ensures that the circle will not drop below the ideal number (8-10) while allowing for there to be some drop off in the circles. Therefore, 24 boys will be randomised for each circle. This means that the delivery team will deliver BAM to 540 boys across 9-10 schools over the one-year internal pilot and two-year efficacy study. As such, power calculations were conducted to determine the minimal detectable effect size (or MDES) on the basis of the following inputs and assumptions:

- **Significance level (two-sided)**: **0.05**
- **Power = 0.8**
- **Total sample size = 1,080**
- **Sample size assessed as being upper secondary or tertiary level of need = 918**
- **Pre-test post-test correlation = 0.5**[28]

Note that since covariates are not included in our power calculations, these are conservative estimates of the achievable MDES

Starting from a maximum possible sample of 1,080 CYP, it is estimated this will result in 918 young people assessed as being upper secondary or tertiary level of need and therefore eligible to be part of our primary outcome analysis sample. This gives us a starting, in principle MDES of **d=0.17**, which aligns with the approximate MDES required by YEF's Magnifying Glass Guidance for the primary outcome.

This will differ from the true MDES associated with our primary outcome analysis, given that some proportion of our sample of 918 CYP may withdraw from the study or fail to provide post-test outcome data. While the operational target being tracked via progression criteria

[28] Goodman (2001)'s original psychometric validation study (quoted in YEF's SDQ guidance) found pre-post correlations of 0.62 in the total score of the self-report SDQ over 4-6 month intervals. It is expected that the pre-post correlation to be slightly lower than this for the longer time interval here, and have chosen 0.50 to be conservative.

and the referral and recruitment tables is 10%, using a more conservative assumption that study attrition may be as large as 25%, this would result in a final analysis sample size of **689**, and mean the study is well powered to detect an effect of **d=0.19.**

While the current proposal suggests to use YEF's levels of need as the tool for defining the sub-sample, this will be reviewed before publication of the SAP. Analysis on baseline data will be conducted to determine how the sample breaks down by YEF levels of need and HSA tiers to determine the estimated sample size.

Table 5: Sample size calculations

		PARAMETER	
Minimum Detectable Effect Size (MDES)		Per protocol, at randomisation: d=0.17	MDES accounting for expected **attrition**: d=.19
Pre-test/ post-test correlations	level 1 (participant)	0.5[29]	
	level 2 (cluster)	NA	
Intracluster correlations (ICCs)[30]	level 1 (participant)	NA	
	level 2 (cluster)	NA	
Alpha		0.05	
Power		0.8	
One-sided or two-sided?		Two-sided	
Number of participants	Intervention	**At randomisation**: 459	Post-attrition: 344

[29] Goodman (2001)'s original psychometric validation study (quoted in YEF's SDQ guidance) found pre-post correlations of 0.62 in the total score of the self-report SDQ over 4-6 month intervals. We would expect the pre-post correlation to be slightly lower than this for the longer time interval here, and have chosen 0.50 to be conservative.

[30] Note that circle level clustering and the impact on power will be explored in the Statistical Analysis Plan.

	Control	**At randomisation**: 459	Post-attrition: 344
	Total	**At randomisation**: 918	Post-attrition: 689

A post-attrition MDES is provided in the table above given that statistical power depends on the observations available for the primary outcome analysis. Incorporating an anticipated attrition rate ensures that, after expected loss to sample size, the study retains sufficient power to detect the prespecified effect size at the chosen significance level. This approach is intended to reduce the risk of Type II error due to loss of statistical power and does not mitigate any potential bias introduced by differential attrition, which will be addressed separately in the analysis plan. Because missing data can occur through attrition, recording errors, or matching failures, this will be systematically addressed by documenting missingness rates for all key variables and outcomes across trial arms. If missing data exceeds 5% overall or differs by more than 5 percentage points between groups, the underlying mechanisms will be investigated and appropriate statistical corrections applied. This investigation will include comparing baseline balance tables for both the originally randomised cohort and the post-attrition sample to monitor for any deterioration in equivalence. Furthermore, statistical tests will be conducted for both differential and selective attrition to identify if specific participant profiles dropped out systematically. These procedures aim to ensure that any potential biases introduced by missing observations are transparently assessed and accounted for in the final evaluation.

Impact Evaluation: Methods and data collection

Data sources

Our main sources of outcome data are:

1) Young person outcome surveys: Young people will be invited to complete outcome surveys at three time points - prior to randomisation (baseline) and immediately after the intervention concludes (endline) and at 4-month follow up (approximately 13-months post-randomisation). Surveys will be administered in person in a school room using tablets or computers with the support of an evaluation team member. The survey will be done in groups tailored to the needs of the participants and scheduled at a time the school recommends to limit disruption to the school day. The survey will capture intermediate and short-term outcomes as anticipated in the ToC first, such as self-esteem, emotional regulation, and social skills, as well as medium-term behavioural outcomes such as conduct and prosocial behaviour.

Surveys for all time points will be completed in groups with Lab researchers present. Teachers will advise on student needs in advance so that the experience can be tailored to them (for example, individual vs. small group data collection, support with reading the questions, completion of the questionnaire privately or spoken through with the evaluator, need for breaks). Research team members administering the measures will be experienced in administering surveys with this population. Our aim is for a designated school staff member to be on standby to support with any immediate difficulties, and urgent safeguarding issues will be escalated in line with the school's safeguarding procedures. Young people will be offered a break if needed before returning to class.

All proposed full-scale measures are validated for use in this population and have been used in evaluations of interventions for similar difficulties previously. In selecting the measures to collect it was considered: the balance within the survey of positive and negative questions, a completion time of 15-20 minutes and REDI factors, which have been considered by our expert advisors. The survey length will be monitored and seek input from children and young people on length prior to the start of the evaluation. A group of young people will also be asked to review some survey questions and data collection strategies. All young people (treatment and control) will be offered a £20 voucher in recognition of the time taken completing the post-programme and 4-month follow-up surveys. There is no incentive for completing the baseline survey as this is a required step prior to randomisation and inclusion in the trial.

The Lab recognises that the population of young people targeted by BAM may be more likely than others to be absent from school. The school will be asked to support facilitating survey data collection for young people who are absent from school on the day of data

collection, and/or offer a video call to support young people to complete the surveys. The Lab appreciates this is an ask on teachers' time and therefore a 'mop-up' day of data collection for schools will be offered where there are lots of absent students.

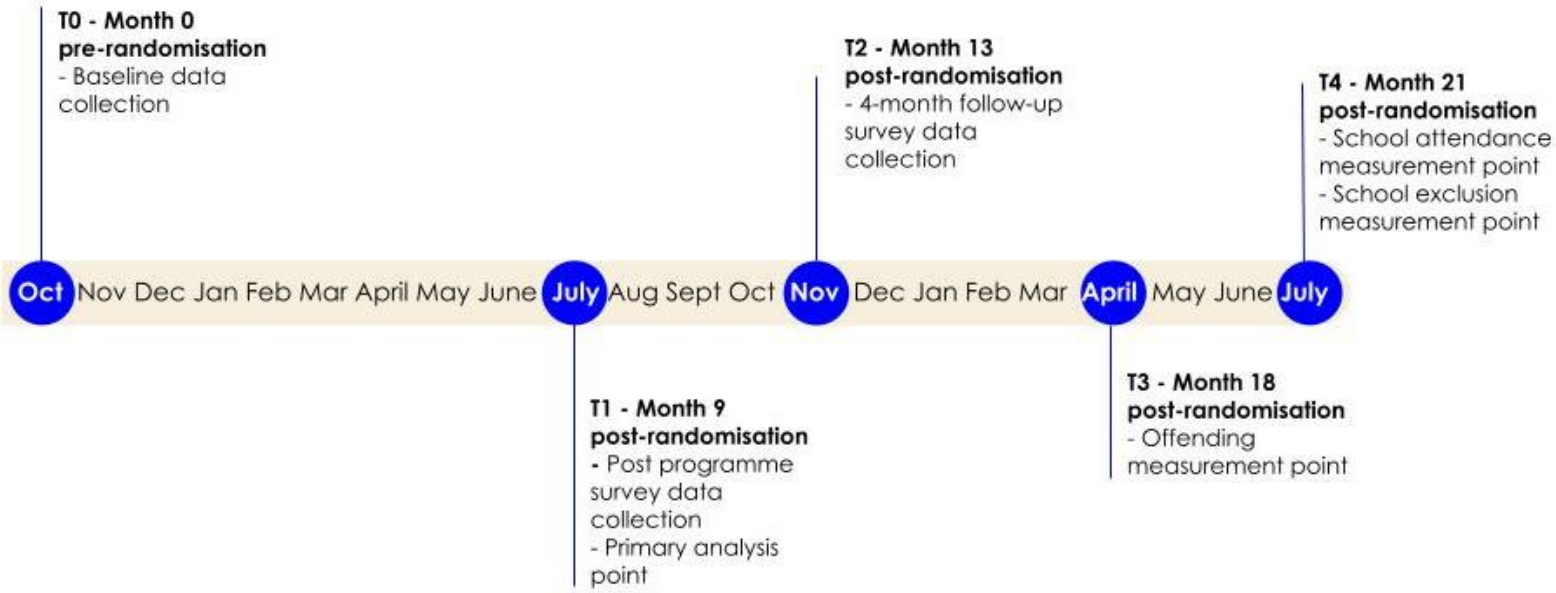
2) **Police Data:** To assess offending a) within a shorter timeframe than is possible through the Police National Computer (where there is a longer-time lag between offence and data availability), and b) more objectively than is possible through self-report surveys, the Lab will attempt to link the trial data with local police data. Arrests are the proposed proxy for offending. Arrests have been chosen rather than other variables included in local police data (such as police outcomes, including cautions) as it is expected that these will be entered into the local police systems at or near the time of the event, prior to any subsequent police outcome, and therefore to serve as a more reliable proxy of offending behaviour within the trial's timeframe. Collecting arrest data covering the trial period and up to 6-months after post-test and the completion of the programme (i.e. approximately 18-months after randomisation) is proposed. An immediate post-programme assessment of local police data might primarily capture novelty effects or the fact that BAM has been occupying children and young people's time, rather than meaningful or sustained change. A 6-month delay may also be better aligned with cycles of police reporting and increase the likelihood that any newer arrests are captured in our data. While we anticipate challenges in obtaining police data, the Lab is already in the process of accessing Met data for other projects. Additionally, having 18 months post randomisation provides us with a substantial period to resolve any delays and issues. The Lab recognises that the use of police arrest data as an outcome requires a nuanced approach, particularly given the well-documented racial disproportionality within the criminal justice system. Evidence highlights that Black boys are significantly overrepresented in arrest statistics relative to the general population, a disparity often driven by systemic factors—such as over-policing and institutional racism—rather than differences in actual offending behaviour alone[31]. We therefore acknowledge that arrest data is not a pure measure of behaviour and is subject to well-documented racial disproportionality in policing. Therefore, we do not view this as an 'objective' truth. Instead, we use it as a necessary triangulation point to complement self-reported survey data, which is itself susceptible to social desirability bias and under-reporting in this cohort. We will interpret findings from this data with significant caution, triangulating them alongside self-reported behavioural outcomes and school exclusions.

3) **Administrative Data:** Administrative data from schools will supplement survey data and provide objective indicators of attendance and long-term outcomes such as exclusions, suspensions and transfers. Based on the Lab's experience and on Islington's understanding of their schools this data is stored in accessible electronic systems and that it would be a

[31] https://youthendowmentfund.org.uk/wp-content/uploads/2025/02/YEF_Racial_Disproportionality_FINAL.pdf

simple task to retrieve this data for the pupils referred to the programme. An excel sheet will be shared between the Lab and each school, held on a secure server, that they can add administrative data to.

Figure 3. Research timeline.



Baseline measures

At baseline, measures will be collected for each outcome variable prior to randomisation. In addition, the referral form will be used to gather data on student demographics (including gender, age, race/ethnicity, and disability), socio-economic status (as indicated by eligibility for free school meals), history of involvement with the criminal justice system, and family care experience. Perceived discrimination (Everyday Discrimination Scale)[32] will also be captured during the baseline survey. Experiences of discrimination in particular groups might mediate their interactions with school-sanctioned support and conflict resolution provisions, classes, and mental health provisions. It might also moderate their outcomes, particularly relating to administrative data. This provides a rich data source that self-reported ethnicity alone cannot capture.

[32]Williams, D.R., Yu, Y., Jackson, J.S., and Anderson, N.B. Racial Differences in Physical and Mental Health: Socioeconomic Status, Stress, and Discrimination. Journal of Health Psychology. 1997; 2(3):335-351.

Outcome measures

In accordance with the theory of change, short- and intermediate psychological and relational outcomes as well as behavioural change outcomes are included. Our specific outcome measures are detailed in Table 6. The Lab acknowledges that participants from different backgrounds might respond differently to these measures, and that some of them might extenuate existing biases. All the measures selected were chosen with this in mind, balancing these issues with the need to measure outcomes in a validated way. Experiences with the surveys will be explored in the qualitative research.

Primary Outcome

The primary outcome is behavioural difficulties, measured using the *self-reported Strengths and difficulties questionnaire (SDQ)* total score at post-programme only for the boys assessed as being at upper-secondary or tertiary level of need. This reflects an intermediate outcome in the theory of change—capturing behavioural improvements expected to emerge after changes in underlying attitudes and competencies as well as relationships.

Secondary Outcomes

Secondary outcomes span short-,medium, and long-term domains within the logic model:

- **Short and medium term outcomes**: Self-control and emotional regulation (Difficulties in Emotion Regulation Scale), Self-esteem (Rosenberg Self-Esteem Scale), Behavioural difficulties (SDQ -total score and externalising subscale) Social skills (SDQ-, prosocial subscale)
- **Long-term outcomes**: offending (measured via arrests in local police data), school attendance, and school exclusions, suspensions, and transfers from administrative data source (measured via school data).

Since we anticipate that BAM can positively drive outcomes for all boys, not just those at higher levels of need, we include all boys in our secondary outcomes except offending.

Offending data

Individual-level data will be accessed on boys based in London, who are in both the treatment and control group. For these boys, the following information will be requested:

- Number of times each child was arrested
- Date of each arrest
- For each offence linked to an instance of arrest:
- Offence type(s) (e.g. violence against the person, sexual offences, robbery)
- Police recorded outcome(s)(e.g. caution, community resolution, no further action)

- Home Office outcome code (e.g. Outcome 1, 2, 8)

To allow the Met to extract data for individuals involved in the evaluation, the following personal identifiers will be provided:

- first name
- last name
- date of birth
- residential address

Note: this is not the full list of CYP data the Lab will share with the Met. This list may be refined in further conversations with the Met. The aim will be to strike a balance between providing sufficient information for the Met to effectively link CYP data, and minimising the sharing of identifiable participant data.

Exploratory outcomes

Measurement of perceptions towards gender norms and stereotypes

BAM serves as an effective upstream intervention for Violence Against Women and Girls (VAWG) prevention, despite not being a dedicated VAWG initiative. It operates through three core theoretical mechanisms: the explicit "Respect for Womanhood" value that fosters gender-equitable norms, the development of core competencies like challenging toxic masculine scripts and practicing respectful social interactions, and the provision of prosocial male role models who embody accountability and emotional intelligence. Together, these components work to reshape underlying attitudes, improve emotional regulation, and provide a model for healthy masculinity, ultimately helping to disrupt the pathways that lead to VAWG perpetration.

Although explicit VAWG topics (such as domestic abuse, stalking, or coercive control) are not the primary focus of the daily curriculum, combining value internalisation, active skill-building, and positive male modelling creates an indirect pathway. Together, these elements transform underlying gender attitudes and emotional regulation capacities, ultimately disrupting downstream VAWG perpetration at its source.

Given the impact on gender attitudes, we are proposing measuring young people's perceptions of gender norms and stereotypes via the Global Early Adolescent Study Gender Stereotypical Traits Scale.

This will be measured via an optional additional survey conducted during the survey sessions, immediately after completion of the outcome survey. YP will be offered £5 for completing the additional survey. Since this is optional, it will not be a prespecified

outcome, rather we will conduct exploratory analysis. If we find high completion rates during the pilot, we will consider including this as an outcome in the efficacy years.

This will only be collected at baseline and post-programme.

HSA subscales

In addition to our specified secondary outcomes, two separate exploratory analyses will be conducted using the HSA data. This will allow further learning about the impact of BAM on important social-emotional development constructs without increasing the level of data collection burden on students.

First, we will conduct exploratory descriptive analysis on subscales from the HSA that are most closely aligned with the BAM Theory of Change. This will include the following subscales:

- Empathy
- Assertiveness
- Learning Interest
- Optimism
- Critical Thinking

Additionally, we will conduct exploratory analysis to understand the overall level of improvement in the level of social-emotional support needs for young people whose baseline score falls in the highest level of need based on their HSA scores (Tier 3). This is based on the Multi-Tiered System of Support (MTSS)[33] employed by PEAR in the HSA.

[33] National Center on Response to Intervention. (2010). Essential components of RTI: A closer look at response to intervention. American Institutes for Research.

Table 6. Outcome measures.

Type of outcome	Outcome measured	Instrument	Completed by	Number of items	Data collection points (Main analysis point bolded)	Subscales to be used	ToC Element	Scoring	References
Primary	Behavioural difficulties for for boys assessed as being upper secondary or tertiary level of need	Strengths and difficulties questionnaire (SDQ)	CYP self-report	25	T0 - Baseline **T1 - Post programme** T2 - Four month follow-up	All subscales (4) - apart from prosocial subscale - and externalising scale separately	Medium-term outcome (reduced behavioural difficulties and externalising problems). Follows the short-term gains in self-regulation and emotional regulation.	Total difficulties Score: A score from 0-40 is generated by summing scores from all the subscales, except the prosocial subscale.	Goodman, 1997
Secondary	Offending (arrests as a proxy for offending) for boys assessed as being upper secondary or tertiary level of need	Administrative data	Local police data	n.a.	**T3 - 18 months post randomisation**	n.a.	Long-term outcome (reduced offending, violent and non-violent). The final outcome in the chain, following reduced behavioural difficulties and improved	A binary indicator describing whether a child or young person has been arrested for an offence in the 18 months following randomisation (no arrests = 0, one or	n.a.

Type of outcome	Outcome measured	Instrument	Completed by	Number of items	Data collection points (Main analysis point bolded)	Subscales to be used	ToC Element	Scoring	References
							school engagement.	more arrests = 1).	
Secondary	Behavioural difficulties for for all students	Strengths and difficulties questionnaire (SDQ)	CYP self-report	25	T0 - Baseline **T1 - Post programme** T2 - Four month follow-up	All subscales (4) - apart from prosocial subscale - and externalising scale separately	Medium-term outcome (reduced behavioural difficulties and externalising problems). Follows the short-term gains in self-regulation and emotional regulation.	Total difficulties Score: A score from 0-40 is generated by summing scores from all the subscales, except the prosocial subscale.	Goodman, 1997
Secondary	Social skills	Strengths and difficulties questionnaire (SDQ)	CYP self-report	5	T0 - Baseline **T1 - Post programme** T2 - Four month follow-up	"Prosocial scale"	Short-term outcome (improved social skills). Comes from the perspective-taking and introspection	Score of subscale.	Goodman, 1997

Type of outcome	Outcome measured	Instrument	Completed by	Number of items	Data collection points (Main analysis point bolded)	Subscales to be used	ToC Element	Scoring	References
							mechanism (roleplay, group missions that require attention to others).		
Secondary	Externalising behaviours	Strengths and difficulties questionnaire (SDQ)	CYP self-report	10	T0 - Baseline **T1 - Post programme** T2 - Four month follow-up	Externalising problems scale - a combination of the conduct and hyperactivity scales	Medium-term outcome (reduced behavioural difficulties and externalising problems). Isolates conduct and hyperactivity, rather than the combined difficulties score.	Combined score of the two subscales.	Goodman, 1997
Secondary	Self-esteem	Rosenberg Self-Esteem Scale (RSES)	CYP self-report	10	T0 - Baseline **T1 - Post programme**	n.a.	Short-term outcome (improved	Total scores range from 0 to 30. Items	Rosenberg, 1979

Type of outcome	Outcome measured	Instrument	Completed by	Number of items	Data collection points (Main analysis point bolded)	Subscales to be used	ToC Element	Scoring	References
					T2 - Four month follow-up		self-esteem). Comes from the sense of achievement and sense of belonging mechanisms.	are assessed on a 4-point Likert scale. Overall, higher scores indicate higher self-esteem.	
Secondary	Self-control and emotional regulation	Difficulties in Emotion Regulation Scale (DERS)	CYP self-report	36	T0 - Baseline **T1 - Post programme** T2 - Four month follow-up	All subscales (6).	Short-term outcome (improved self-control; improved emotional regulation). Comes from the self-regulation mechanism (learning to hold back, calm spreading through the group) and the reducing-automaticity mechanism (a trained pause	Total scores range from 36 to 180. Items are assessed on a 5-point Likert scale. Overall, higher scores indicate greater difficulties in emotion regulation.	Neumann et al., 2010

Type of outcome	Outcome measured	Instrument	Completed by	Number of items	Data collection points (Main analysis point bolded)	Subscales to be used	ToC Element	Scoring	References
							before reacting).		
Secondary	School attendance	Admin data	School-level administrative data	N/A	**T4 - 21 months post randomisation**	N/A	Medium-term outcome (improved school attendance). Follows the short-term gains in self-regulation, self-esteem, and social skills.	Percentage attendance over the year following the programme	NA
Secondary	Exclusions, suspensions and transfers	Admin data	School-level administrative data	N/A	**T4 - 21 months post randomisation**	N/A	Medium-term outcome (reduced school exclusions, suspensions, transfers). Follows the same short-term mechanisms as attendance.	A measure capturing the likelihood that the CYP faced either exclusion, suspensions or transfers (faced any of the three issues = 1; 0 otherwise). Combined measure as	NA

Type of outcome	Outcome measured	Instrument	Completed by	Number of items	Data collection points (Main analysis point bolded)	Subscales to be used	ToC Element	Scoring	References
								likelihood of any of these events is low.	
Exploratory	Offending (arrests as a proxy for offending) for all boys	Administrative data	Local police data	n.a.	T3 - 18 months post randomisation	n.a.	Long-term outcome (reduced offending). Same element as the primary offending outcome above, measured here across the full cohort.	A binary indicator describing whether a child or young person has been arrested for an offence in the 18 months following randomisation (no arrests = 0, one or more arrests = 1).	n.a.
Exploratory	Perceptions of gender norms and stereotypes	Global Early Adolescent Study Gender Stereotypical Traits Scale	CYP self-report	7	T0 - Baseline T1 - Post programme	n.a	Short-term outcome (Improved equal attitudes towards gender)	Total score	Moreau et al., 2021

Type of outcome	Outcome measured	Instrument	Completed by	Number of items	Data collection points (Main analysis point bolded)	Subscales to be used	ToC Element	Scoring	References
Exploratory	Empathy	Holistic Student Assessment Subscale	CYP self-report	n.a	T0 - Baseline T1 - Post programme T2 - Four month follow-up	n.a	Empathy	Total Score	PEAR, 2011
Exploratory	Assertiveness	Holistic Student Assessment Subscale	CYP self-report	n.a	T0 - Baseline T1 - Post programme T2 - Four month follow-up	n.a	Assertiveness	Total Score	PEAR, 2011
Exploratory	Learning Interest	Holistic Student Assessment Subscale	CYP self-report	n.a	T0 - Baseline T1 - Post programme T2 - Four month follow-up	n.a	Learning Interest	Total Score	PEAR, 2011
Exploratory	Optimism	Holistic Student Assessment Subscale	CYP self-report	n.a	T0 - Baseline T1 - Post programme T2 - Four month follow-up	n.a	Optimism	Total Score	PEAR, 2011
Exploratory	Critical Thinking	Holistic Student	CYP self-report	n.a	T0 - Baseline T1 - Post programme T2 - Four	n.a	Critical Thinking	Total Score	PEAR, 2011

Type of outcome	Outcome measured	Instrument	Completed by	Number of items	Data collection points (Main analysis point bolded)	Subscales to be used	ToC Element	Scoring	References
		Assessment Subscale			month follow-up				

Analysis

Main effect outcome analysis

The analysis will follow a pre-specified Statistical Analysis Plan (SAP) to be finalised prior to the commencement of data analysis. The approach will differ for the pilot and efficacy phases of the trial.

For the pilot phase, the analysis will be primarily descriptive. The focus will be on summarising key progression criteria, including recruitment rates, data completeness, and participant retention. The pilot trial will therefore inform the decision on whether to proceed to the full efficacy trial, as well as provide more reliable inputs for estimating effect sizes.

If the project progresses to efficacy, outcome data from both the pilot and efficacy phases will be analysed together to determine the impact of the programme. All outcome data will be analysed using an Intention-to-Treat (ITT) analysis, including all randomised participants in their assigned groups regardless of the level of intervention received. Regressions will be comparing those in the control arm (or BAU) to those in the treatment arm (or BAM). Analysts won't be blind to treatment allocation as allocation knowledge is necessary for effective analysis to be done. All regression results will be supported by IPE findings where possible.

In all regression models, we will adjust for baseline pupil-level characteristics, including pre-intervention outcomes, ethnicity, age, perceived discrimination, and FSM eligibility. To account for the multi-site structure of the trial—where recruitment, BAU provision, and randomisation occur within schools—our regression models will include school fixed effects (γ_j). This controls for unobserved school-level factors, effectively comparing treatment and control pupils within the same school environment.

The mathematical specification for the regression model for pupil *i* in school *j* is:

$$Y_{ij} = \alpha + \beta_1 * Treatment_{ij} + \beta_2 * Control_vector_{ij} + \gamma_j + \varepsilon_{ij}$$

Where Y_{ij} is the outcome for pupil *i* in school *j*; $Treatment_{ij}$ is the indicator for the treatment arm; $Control_vector_{ij}$ represents pupil-level covariates; γ_j represents fixed effects for school *j*; and ε_{ij} is the residual error term. To account for the dependency of observations within schools, standard errors will be clustered at the school level.

Where: Y_{ij} is the measure of the outcome at the primary analysis point as per Table 2; $Treatment_i$ is the indicator variable for the treatment arm; $Control\ vector_i$ represents the

vector of fixed effects for all pupil-level covariates; and ε_i is the residual error term. Given our sample size is sufficiently large, Cohen's d and Hedges' g are numerically almost identical (correction factor ≈ 0.999). Therefore, for continuous variables, Cohen's d is reported, which is equivalent to Hedges' g in this context. For binary administrative outcomes, such as arrests, we will report the Odds Ratios (OR) derived from the logistic regression, alongside calculated Risk Ratios (RR) and risk differences to provide highly interpretable metrics of programmatic impact.

Our primary analysis will consist of analysing the SDQ on the subgroup of students who are assessed as being at upper-secondary or tertiary level of need. Conduct analysis will only be conducted of offending behaviour on this sub-sample as these are the young people who will be most vulnerable to involvement to violence and most likely to demonstrate significant improvements in our outcomes if the intervention is effective. Secondary analysis of the SDQ will be conducted on all boys, but will only report exploratory analysis of offending for all participants and provide commentary on it, since the trial is not powered to find an effect for this cohort.

Subgroup analysis

To explore differential impacts of the BAM programme, we will conduct pre-specified exploratory subgroup analyses based on ethnicity, free school meal (FSM) status, and baseline perceived discrimination scores. Our assessment of differential impact will be conducted using treatment-by-subgroup interaction models integrated into the primary regression framework. For each subgroup, an interaction term (Treatment × Subgroup) will be added to the regression model. A statistically significant coefficient on this interaction term will provide formal evidence of a differential treatment effect between groups.

In addition to these interaction models, we will report descriptive subgroup-specific treatment estimates (mean differences) with their associated 95% confidence intervals. This dual approach ensures that we can both formally test for differential impact and present the raw magnitude of the effect within each specific stratum.

All subgroup analyses are considered exploratory. Given the risk of Type I error associated with multiple testing, results will be interpreted with caution, and we will not perform multiple comparison adjustments. We will report the p-values for all interaction terms and provide a narrative discussion of any identified trends, triangulating these quantitative results with qualitative findings from the Implementation and Process Evaluation (IPE) to provide context.

Implementation and process evaluation

Research questions

Table 7. IPE research questions.

Establishing Impact		RQ 1.1: What is the perceived impact of the programme?
Establishing mechanisms and moderating factors		RQ 2.2: What are the perceived mechanisms creating the programme's impact?
Assessing acceptability	RQ 3: Is the programme acceptable to stakeholders?	RQ 3.1: Is BAM acceptable to young people and schools? If not, why not? RQ 3.2: Are there any barriers/facilitators of programme acceptability? Are these the same/similar to those expressed during the pilot trial? RQ 3.3: Do views on acceptability of the intervention differ amongst different groups of young people, with a specific focus on young people from racially minoritised backgrounds. RQ 3.4: Does acceptability vary depending on how the programme is delivered?
Assessing deliverability	RQ 4: Is the programme being delivered with fidelity and what factors impact fidelity?	RQ 4.1: Is the programme being delivered with fidelity? If not, why not? RQ 4.2: What are the barriers and facilitators

		to delivering the programme well and with fidelity? RQ 4.3: What variations in delivery are appropriate for effective implementation? RQ 4.4: What variations in delivery are not appropriate for effective implementation? RQ 4.5: Does the level of fidelity differ across different practitioners? RQ 4.6: Does the level of fidelity differ for those schools that have been included in the pilot and efficacy trial compared to those that are only involved in the efficacy phase? RQ 4.7: Does the level of fidelity vary in the schools who have already delivered BAM?

Research methods

A mixed methods approach to data collection is adopted for the IPE with themes explored through both qualitative and quantitative data collection methods. Findings will be triangulated across methods to answer research questions.

Data from the following sources will be triangulated:

Administrative/monitoring data

Essential administrative programme data will be collected from schools and counsellors. Data on referrals and attendance will be captured to monitor programme recruitment, retention and reach. Data on programme delivery (e.g. activities completed in each school, logged by counsellors) will be used to monitor dosage and fidelity. Methods for capturing data on fidelity will be refined during the co-design stage and building on existing data collection mechanisms.

This data will also be used to inform the sampling approach for the qualitative work.

Surveys

All counsellors and young people taking part in the programme will be invited to take part in short, online surveys to capture data on key aspects of deliverability, acceptability and evaluability. The following surveys will be conducted:

1. **Young person satisfaction surveys:** A short survey will be distributed to CYP in the treatment arm in the final BAM session at the end of the term of delivery. The survey will be short and accessible to ensure active engagement from CYP from diverse backgrounds. CYP will be able to complete surveys independently or with support from the practitioner or a school staff member. CYP will be assured during the process that their responses will remain confidential, and opinions will not be attributed to any individuals. The survey will be piloted with the YAG during the mobilisation phase to gain feedback on aspects of survey design, e.g. survey length, question design, question language and question content and the survey will be refined based on this feedback. The survey will measure CYP views on fidelity components, expectations and feedback on the intervention, quality of support received[34], and overall acceptability of the intervention and trial.

[34] Measured using 5 items from the Mentor-Youth Alliance scale

2. **Counsellor satisfaction surveys:** Counsellors will be invited to complete an online survey at the end of each cohort during the pilot and efficacy trials to align with the focus groups. The survey will focus on counsellors' experience of delivery, acceptability of the programme, adaptations made and counsellor wellbeing (a key potential backfire effect).
3. **Outcome surveys:** Endline data collection with young people and school staff will contain a small number of questions to gather broad feedback from the whole sample on key IPE questions, including outcomes, mechanisms, backfires, and perceptions of the control arm offer (for those in the business as usual arm).

Quality of delivery, fidelity to the delivery model and trial arrangements will be examined by using information from the surveys and administrative data including the fidelity tracker. Data collected through these activities will inform the issues to explore in the qualitative component (e.g. reasons for low fidelity, different levels of CYP engagement).

Qualitative research

In-depth interviews will be primarily used with counsellors, school staff, young people taking part in the programme and their caregivers, as these are responsive, flexible and allow for detailed exploration of views and experiences. Paired interviews/mini group discussions will be used where it is feasible and valuable for gathering different perspectives in a single encounter. Qualitative data collection will be guided by topic guides which will list the key topics, themes and prompts to be covered. With participants permission, interviews will be audio recorded and transcribed verbatim to ensure accuracy and facilitate subsequent analysis. Wherever possible to minimise disruption, these interviews will be conducted in conjunction with the endline data collection visits. The boys will be purposively sampled within each cohort to maximise diversity of views, including engagement with the programme, ethnicity and gender. Interviews with boys, counsellors, school staff and parents/caregivers will be conducted in person where possible during a visit to the school but telephone or online platform interviews will also be offered for those not available during the visit.

In-depth interviews/mini group discussions will be conducted with counsellors and school staff to explore feasibility and acceptability of the trial, recruitment processes, feedback on the training and support received, acceptability and feasibility of the practice model, young people's engagement, implementation barriers and facilitators, and perceived impacts. Bias will be minimised in this sampling by random selection of participants. Interview/ group discussion questions about the perceived impact and mechanism will help explore and validate the ToC. Topic guides will build off questions used during the BAM Qualitative study (Chapin Hall 2016) which could have a confirmatory role in the exploration of whether the Theory of Change and BAM Mechanisms are supported through this adapted approach.

Interviews will also be carried out with:

1. Young people taking part in the programme (target of 30 across the 3 delivery cycles, comprising approximately 10 control YP and 20 treatment YP; sampled at a rate of 1-2 per selected circle) to explore their experience of the programme and its core components (e.g. appropriateness of the number of sessions), other services received, perceived impacts, and experiences of trial procedures (including outcomes measured). It will specifically be explored whether acceptability and perceptions of quality differ for a) young people from different backgrounds (e.g. racial, cultural, ethnic, class, etc), b) BAM groups with higher levels of need/risk of violence. Various characteristics will be monitored to identify where our sample will allow for intersectionality analysis.
2. Parents/caregivers of young people taking part in the programme (up to 10 across the 3 delivery cycles) to explore acceptability of randomisation, the best way to obtain consent, factors considered when providing parental consent, and views on CYP engagement with the intervention.

Trained researchers will collect qualitative data. The interviewers will be from BIT's qualitative research team and will all have significant experience conducting interviews with vulnerable participants. However, given the sensitivity of the project and the focus on participants from marginalised backgrounds, these researchers will receive specific input from our race equity partner ClearView and be guided by ClearView on how to build authentic rapport and navigate power dynamics to ensure cultural competence and safety in this specific context. To further support young people to feel comfortable in the research space, creative, trauma-informed tactics will be employed instead of standard formats. This includes offering fidget toys to aid regulation, incorporating drawing or writing tasks to allow for non-verbal expression, and ensuring flexibility for breaks throughout the session. A 'traffic light' system will also be employed, using coloured cards, to give participants a physical, non-verbal way to pause or skip questions, giving them tangible control over the interview process.

The YAG and our race equity partner ClearView will be consulted regarding the content of the topic guides. This ensures that lines of inquiry are culturally sensitive and do not inadvertently reinforce deficit narratives. Specifically, input will be sought on how to frame questions regarding adversity to ensure participants feel safe discussing systemic barriers to engagement with the intervention, rather than focusing solely on individual behavioural traits. ClearView will train researchers on how to respond to social norms and power dynamics as much as possible.

Table 8 below summarises each of the data collection methods for the IPE and the research questions they will explore.

Table 8. IPE data collection methods and addressed research questions.

Research methods	Data collection methods	Participants/ data sources (type, number)	Research questions addressed
Qualitative data	Interview	Counsellors and school staff (n = 10-15 across 3 delivery cycles)	Evaluability Acceptability
		Young person interview (n = 30 across 3 delivery cycles)	Evaluability Acceptability
		Parent/care givers interview (n = 10 across 3 delivery cycles)	Evaluability Acceptability
Qualitative and quantitative data	Survey data	Young person satisfaction survey	Impact Outcomes Acceptability
		Counsellor satisfaction survey	Impact Outcomes Acceptability
Quantitative data	Administrative data	Data on referrals and attendance	Evaluability Establishing mechanisms and moderating factors
		Data on programme delivery	Evaluability Establishing mechanisms and moderating factors Deliverability

Fidelity

While the exact measure of fidelity is to be worked out dependent on the final adaptations, it is expected that it follows a similar process to what was used in the feasibility and pilot studies, which is in line with Youth Guidance's robust approach to fidelity monitoring. Fidelity will be assessed by Youth Guidance.

Below is a summary of the approach expected to form the basis for the fidelity measure, based on what was used in the previous studies.

1. Self-Reported Adherence (Implementation Checklists)

- **Method:** Counsellors are required to document the details of all services provided to participants, including **implementation checklists** after every *BAM Circle* session and any interactions with any YP (not just treatment YP) outside of BAM circles.
- **What is measured:** These checklists record which specific activities are delivered compared to those specified in the lesson plan. There is also the opportunity to record any adaptation made during the session along with the reason that an adjustment was made.
- **Metric:** "Adherence" is calculated based on the completeness of these activities; greater completeness indicates higher adherence to the manual. Adaptations coupled with the rationale for adaptation (required due to immediate context vs. unplanned adjustment) is also recorded.

2. Expert Observation (Replication Specialist Assessment)

- **Method:** A "Replication Specialist" from Youth Guidance (the US developer) conducts formal assessments of the counsellors.
- **Frequency:** Observations typically occur during site visits (e.g., five visits per year in the pilot) or via video review when in-person visits are restricted (e.g., during COVID-19).
- **What is measured:** The specialist assesses counsellors against **five core competencies**:
    1. Systemic leadership (engaging school staff/parents)
    2. Clinical skills (listening, assessing, challenging)
    3. Modelling (demonstrating values)
    4. Group work
    5. Youth engagement
- **Scoring:** Counsellors are rated on a five-point scale ranging from "Basic Knowledge" to "Expert". These ratings are made early in the program year as baseline, and at the end of the program year to determine growth in these competencies based on training and coaching received.

3. Curriculum Progression (Dosage/Reach)

- **Method:** Routine monitoring data tracks how far each group progressed through the *BAMual* (curriculum) along with how consistently the sessions are delivered and attendance rates of the participants.

- **Target:** The target is for counsellors to deliver all curriculum lessons over the delivery period of 31 weeks. Progress is tracked by noting the "lesson reached" for each group at the end of each year. Each Lesson is designed to be completed in one session. There may be some sessions where the Counsellor may need to address an incident that happened with the young people or in the school/community ("Response to Incident"). Schools may occasionally need sessions to be cancelled. With these considerations, the 25-lesson BAM 1-year curriculum should be able to be completed. The average group attendance rate will be 70%.

4. Fidelity Targets

Fidelity targets for the 1-year programme will be finalised once the programme is complete. However, to provide some context, the feasibility study set specific success criteria for fidelity to determine if the programme was ready for a pilot. These were

- **Target:** Counsellors deliver all lessons from the 25-lesson manual.
- **Target:** Counsellors deliver 27 *BAM Circle* sessions over the single school year.

5. School Enabling Context

- **Method:** The programme Supervisor, in partnership with the BAM Counselor, conducts a "Conditions for Success" assessment to determine the degree to which the school's enabling context facilitates or hinders quality implementation of BAM.
- **Frequency:** Assessments are completed early in the school year to identify what conditions are in place and what needs to be addressed. The assessment is repeated at the end of the school year to determine progress and degree of enabling context.
- **What is measured:** There are five Conditions for Success on which each school is assessed:
    a. School Administration Support
    b. Teacher Support
    c. Adequate Space
    d. BAM Staff Presence in the School
    e. Communication with Data
- **Scoring:** Each dimension includes three to four questions, each rated on a 3-point Likert scale. An Overall Score is calculated as the average across all questions and schools are rated High, Medium, or Low level of enabling context based on the Overall score. If a school scores low on having an enabling school context and the issues are identified early in the year or arise during the school year, the implementation team uses continuous improvement approaches to address the issues and repeat the assessment at an appropriate time. If the issues are identified later in the year, discussions with the school leadership are scheduled, during which

the School level agreement is reviewed. If the issue is not addressed within the time frame agreed, an alternate school may be included in the trial.

Analysis

Data analysis

Data from each element of the IPE will be analysed separately, then triangulated and integrated, identifying areas of difference and reinforcement, and using different data sources to substantiate and explain findings. This includes data to validate the Theory of Change and the hypothesised mechanisms.

Quantitative IPE data
Quantitative data (survey responses, engagement data etc.) will be reported descriptively. Additionally descriptive statistics will be provided on how these data vary across key demographics (age, ethnicity) but expect the sample size to be too small to make any statistical inference about differences between these groups. The intersections across key demographics will also be reviewed where our sample sizes allow.

Qualitative IPE data
Qualitative data will be analysed using a version of the framework approach, which is widely used in applied social research. Framework analysis will be undertaken to examine and interpret qualitative data, with themes developed both deductively and inductively to include unexpected issues. Framework analysis facilitates systematic theme-based comparison within study populations (e.g. between different groups of CYP) and between study populations (e.g. comparing the views of practitioners, parents, school staff , and CYP), as well as within case analysis (e.g. exploring how the particular school context impacts on feasibility and perceived outcomes).

Insights of our race equity partner ClearView and advisory board will also be utilised when approaching qualitative sub-group analysis, and qualitative analysis more broadly. This will further ensure the positionality of the research team is considered, and an equitable approach to analysis is fostered throughout.
Integrating the YAG, enables members to function as peer researchers. This collaboration supports the research team in identifying biases or hidden subtext. This ensures findings and interpretations findings and interpretations are culturally sensitive, valid, and not pathologising.

Cost data reporting and collecting

The cost of delivering the intervention will be reported in the final report, following YEF costing guidance.

This will involve:

- A bottom-up costing approach and break costs down into: prerequisites, set-up costs, and recurring costs. Cost is primarily expected to be driven by staff costs (cost of practitioner, supervisors, and managers involved in delivering BAM), and programme procurement costs (costs for training staff, programme-specific materials and required licenses).

- The total cost for a typical single school receiving the intervention for one round of delivery and the costs per participant for one round of delivery, assuming full compliance, broken down into prerequisites, set-up costs, and recurring costs. The cost estimates will not include costs of the evaluation and the costs of the programme adaptation.
- Depending on heterogeneity in costs across schools, either average costs will be reported, or select a case which is regarded as most representative of the costs expected to be incurred in future, typical rounds of delivery.

Islington is the key delivery partner for this evaluation however schools also incur costs to provide the intervention which would be covered under YEF's costing guidance. To report cost at the end of this study, a template will be produced for the delivery team and a sample of 6 schools to complete, covering staff costs, equipment/materials costs, programme procurement costs, and buildings and facilities costs. This information will be collected from Islington at the end of each term of delivery, across both pilot and efficacy phases and from schools at the end of their term of delivery.

Diversity, equity and inclusion

This section outlines the steps that will be taken to ensure the trial is accessible to all, welcoming, and inclusive. YEF have specified that they expect at least 40 percent of participants will be from racially minoritised backgrounds, specifically those from Black (African and Caribbean), Asian, or other ethnic minority groups. These targets are vital to ensuring that the work genuinely reflects and responds to the communities most affected by serious youth violence. This target is slightly higher than most YEF trials which typically set a target of 30 percent, and partly reflects BAM's origin in working with minoritised communities, including in the US evaluations which served populations who were almost exclusively black (approximately 70 percent) and hispanic (approximately 30 percent). However, ensuring that the trial is designed to engage young people from diverse backgrounds has been central to our co-design process. Our design aims to minimise barriers

to engagement in research activities for minoritised groups and those who are disproportionately affected and/or underserved by existing services or delivery:

1. **Co-design considerations:** Considering DEI throughout the co-design phase by (i) working closely with YG, MHF and Islington's team who have collective deep experience related to delivering programming for young people from Black and minority ethnic backgrounds and (ii) involving YEF's Race Equity Associate, including having them attend meetings and co-design workshops and reviewing this protocol.
2. **Race Equity Partner involvement:** Involving our race equity partner ClearView, experts research consultants specialising on the recruitment and engagement of ethnically diverse research participants, throughout the evaluation design and delivery. This will involve reviewing the protocol, supporting in preparing participant facing materials, and being involved in key decisions around evaluation considerations.
3. **Youth advisory group:** Creating and engaging a youth advisory group (YAG) with a diverse set of young people who have recently completed BAM to inform our approach and discuss key issues. The YAG will be recruited by ClearView. The YAG will be involved in a number of key stages throughout the evaluation process, including in the finalisation of the evaluation design, reviewing participant facing materials, discussing our analysis and synthesising findings.
4. **Sampling of research participants:** Exploring the possibility of sampling research participants based on ethnicity to ensure the research sample is representative of the communities BAM engages with and a range of perspectives and backgrounds are captured. It will be ensured that there is space within interviews and focus groups to explore specific equity and diversity issues.
5. **Contextualising difficulties:** BAM is structured to avoid pathologising the young person's difficulties, instead situating challenges within the context of structural and environmental realities, such as experiences of discrimination and socioeconomic adversity. BAM is a strength-based programme and the sessions are both challenging and reaffirming.
6. **Targeting high-risk populations:** While the BAM programme does involve a variety of participants, the focus is predominantly on young people exhibiting difficulties with emotional regulation and behavioural issues, ensuring resources are directed toward those most likely to benefit. Where possible, schools will be recruited that are in areas of high socioeconomic deprivation or have a high proportion of racially minoritised students.
7. **Accessible Language and Materials:** All information documentation provided to young people will be accessible and written in plain English. Producing a video version with someone explaining the materials in plain English will be explored. Parental consent and information sheets are written in clear, understandable language. All

written material will be reviewed by the YAG to ensure clarity and non-stigmatisation. For young people, an additional Easy Read version of the information sheet will be produced. Information sheets will contain a QR code, phone number and email address so parents/carers, young people and school staff can speak about the trial with the evaluation team prior to giving consent in order that they can ask any questions or explore any concerns prior to giving informed consent.

8. **Accessible data collection:** Striving for equality of access by enabling online (remote) participation in interviews and offering different locations and times for interviews to facilitate access.
9. **Accessible, inclusive and culturally sensitive material and survey design**: Broad survey content will be reviewed by ClearView, and will also be reviewed by our YAG prior to the pilot phase beginning. Multiple tactics will be used to ensure that data collection for young people and school staff is flexible and accessible such as the choice of having questions read aloud and offering breaks (see data collection section for further details).
10. **Compensation and Incentives:** Young people will be compensated for their time and expertise in completing surveys and engaging in interviews, promoting fairness for their contribution to the research.
11. **Minimising Power Imbalance:** In qualitative data collection, an approach will be used which includes tools and techniques to minimise the power imbalance between researchers and participants, such as ensuring inclusive facilitation through displaying empathy and respect. Consideration is given to aligning researcher and participant identities (or utilising peer researchers) where feasible, to improve the perceived safety and trust.
12. **Ensuring best practice:** All researchers conducting interviews will have undertaken NSPCC's Introduction to safeguarding and child protection training and complete a pre-interview training on interviewing best-practice. Class 13's safeguarding training which is grounded in race equity will also be explored.
13. **Measuring discrimination:** The evaluation plans actively seek to measure the direct experience of discrimination, which is recognised as a cumulative risk factor for engagement in violence. For BAM, this involves measuring young people's direct experience of discrimination via the short Everyday Discrimination Scale and including this as a control variable in our analysis.
14. **Contextualised and Sensitive Reporting:** All findings will be interpreted through a collaborative process with our youth advisory group, and our race equity partners ClearView. This will ensure our final report goes beyond simply presenting numbers, by fully accounting for the racialised context of youth justice, explaining data limitations, and communicating findings in a sensitive and impactful way.

15. **Development of a stigma mitigation plan:** The Lab will work with the delivery partners, ClearView, YEF's Race Equity Advisor and our YAG, to ensure a clear stigma mitigation plan including all communications and materials minimise potential stigmatisation. Where necessary, counsellors will work with teachers and school staff to ensure their understanding of the programme minimises stigma.
16. **Reflecting local communities:** In the UK, BAM works with populations reflective of the local communities. It is to be anticipated that the delivery team will work with a portion of young people from minoritised ethnic backgrounds constituting approximately 50-75% of the cohort. Across all schools in the study, proportions of young people from a minority ethnic background ranges from 43.6% to 86.9%.

17. **Contextualising results and mitigating harmful assumptions:** It is crucial that findings are reported with significant care. Specifically, it needs to be ensured that results are contextualised to avoid reinforcing negative racial stereotypes or conflating structural inequities with individual behavioural deficits. Particularly, Counsellors, themselves reflecting the young people they work with in a way that facilitates the building of rapport, work closely with schools to build a positive image of BAM. This image is founded in an elective participation approach and premised on aspiration rather than punishment, and pride in membership rather than shame for behaviour that leads to being 'sent to' BAM. While BAM recognises the potential for harmful and racist stereotypes to manifest in the way that BAM is interpreted by schools, Counsellors will work to forge a positive image for BAM that mitigates some of the assumptions that might be made about the target group.

Young person contribution

The principle of participatory practice is central to our research approach, ensuring the amplification of voices of young people and communities with lived experience of adversity, discrimination, and exclusion. The pilot and efficacy trial will engage young people via a youth advisory group (YAG). This will be held at four key points during the trial with a group of young people that represent the target group for BAM. School staff will also be engaged with during the set-up phase.

Advisory groups will be utilised as active participants in the research journey, recognising that it is easy to only include stakeholder input as input after research decisions have been made. Instead, the advisory group will be involved throughout the research, inviting them to help in the decision making process around key trial decisions such as how consent is sought and how the evaluation and programme is presented. Feedback will be sought on information sheets and other materials and discuss and in any solution-generation around any issues with delivery or evaluation that occur during the trial. Finally the interpretations of the findings will be presented, requesting their input and feedback on the interpretations

assisting in interpreting findings, and helping to shape final recommendations.
Our approach will involve engaging the Youth Advisory Group (ages 11 - 16) recruited by ClearView, who will meet regularly with members of the evaluation team. This approach is designed to allow young people and those with an ethnic minority background to be more active participants in research by embedding the lived experience of adversity, discrimination, and exclusion into the evaluation architecture.
The social context and respecting the boundaries of participants will be considered throughout, ensuring the psychological safety essential for discussing sensitive topics is maintained alongside flexible contribution. Options for how those within the advisory boards can contribute will be provided, ensuring involvement is meaningful and shaped by individual preference rather than predecided by researchers. This approach considers power dynamics and accessibility and ensures the benefits of involvement are defined by the participants themselves. Crucially, experts by experience will be fairly compensated, acknowledging systemic inequalities and the current cost of living crisis.
An example of how experts by experience could be involved includes offering opportunities for co-designing meeting agendas which ensures that issues prioritised by minoritised communities are central to the evaluation, rather than peripheral. Similarly, allowing members to select their preferred engagement methods (e.g., virtual vs. in-person, written vs. verbal) respects the diverse logistical realities of families from marginalised backgrounds.

Recruitment diversity

Schools will put forward young people for the programme that they feel would most benefit. The Lab recognises that there is a risk of this perpetuating selection biases, for example young people from certain subgroups being less likely to be put forward as school has 'given up' on them, or behaviour being interpreted differently through a racialised lens. In both the pilot and efficacy trials, characteristics of young people referred to the programme, consented to the programme and disengaged from the programme/trial. It will be evaluated whether characteristics of these groups (such as gender, age, ethnicity, FSM status and additional educational needs) are reflective of the wider student body. If not, further information on this issue will be sought, via IPE interviews, administrative data or informal conversations with school, practitioners, parents and young people to try and understand the reason for this, and to adjust the approach to the evaluation or programme delivery to better meet the needs of minoritised young people.

Due to the elective nature of BAM participation the Lab recognises that an element of self-referral will contribute to the building of the groups. While this will not entirely erase the risk of selection bias, the groups are finalised through cooperation between the school pastoral team, special educational needs coordinator, BAM Counsellor and other school staff, as well as direct contact with young people, thus allowing Counsellors to make decisions based on the profile of young people as presented through a range of lenses.

Evaluation team experience

The evaluation team has extensive experience working with vulnerable and marginalised communities. The Lab has considerable expertise in violence reduction interventions and evaluating psychologically informed programmes with vulnerable groups. Our team members have been involved in multiple projects focused on working with marginalised communities including the YEF sponsored Generation PMTO study (Tom McBride, Lilli Wagstaff, Emma Forsyth), an evaluation of a YEF funded Boxing intervention for young people at risk of involvement in violence (Tom McBride, Lilli Wagstaff), YEF funded evaluation of Face It (Tom McBride, Lilli Wagstaff), an evaluation of a child-led therapy for children who have experienced domestic abuse (Lilli Wagstaff, Tom McBride, Lisa Mueller).

All staff conducting qualitative interviews will have completed training in inclusive and trauma-informed qualitative research processes, local safeguarding procedures, NSPCC's Introduction to safeguarding and child protection training, and interviewing best-practice. The Lab will also work closely alongside our race equity advisor ClearView, and YEF's Race Equity Associate.

Ethics and registration

This trial is self-assessed as being high risk due to the inclusion of high-risk participants in the form of vulnerable young people. Ethical approval will be sought from both an independent panel of external experts the Lab regularly engages for advice and Islington's internal ethical review process. The panel is composed of academics, research experts and practitioner experts, with experience of working with vulnerable children and experience with safeguarding and child protection.
The independent ethics review committee (ERC) will review the following information:

- Ethical review form.
- All participant-facing materials, including consent forms and information sheets for young people.
- Topic guides.
- Safeguarding and distress protocol.

The ERC will discuss any issues raised by the research with the aim of finding solutions that meet ethical requirements. If there are substantial changes while the research or evaluation is being implemented, the ethics form will be revised and the revisions agreed with the ERC.

The pilot and efficacy trials will be registered at *www.controlled-trials.com*.

Data protection

Appropriate data protection processes will be followed in accordance with BIT processes, including completing a Data Protection and Security Checklist and Data Protection Impact Assessment, which will be reviewed and approved by BIT's legal team.

The Lab will store and handle all data securely and confidentially in line with requirements of the UK GDPR, and Data Protection Act (2018), including that Personal Data shall be processed lawfully, fairly and in a transparent manner that ensures the security of the Personal Data. It is initially proposed that only the Lab and Islington will have access to data collected as part of the evaluation.

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

The legal basis for processing special category data (ethnicity) is Article 9 of UK GDPR, which states that processing special category data is permitted *"for research and statistical purposes which may be in the public interest"*.

The trial will also involve the processing of criminal offence data for the local police offending outcome, including arrest information and associated offence and police outcome data linked to individual trial participants. In additional to the Article 6 lawful basis set out above, this processing falls within Article 10 UK GPDPR. Subject to confirmation by BIT's Legal Team and DPO and agreement of the final data-sharing arrangements with the Metropolitan Police, the Lab intends to rely on Schedule 1 Part 1 Paragraph 4 of the Data Protection Act 2018, which permits processing necessary for scientific or statistical research purposes in the public interest.

The processing is necessary to assess whether BAM affects offending outcomes and will be subject to additional safeguards. Only the minimum identifiable data necessary to enable accurate linkage with police records will be shared; identifiable linkage information will be securely transferred and access restricted; research datasets will be pseudonymised wherever possible; direct identifiers will be separated from analysis data; access to criminal offence data will be limited to authorised members of the research team; and findings will only be reported in anonymised or sufficiently aggregated form. The sharing and processing of police data will also be governed by appropriate data-sharing agreements, retention and deletion requirements, and the safeguards applicable to research processing under the Data Protection act 2018.

The Lab determines that there is a genuine purpose to process this data. This data will inform building the necessary evidence around what works to reduce offending and build young people's social and emotional competencies. Data processing is necessary to complete a robust evaluation. In strict adherence to the principle of purpose limitation, personal data will only be collected and processed for this explicit evaluation of the Becoming a Man (BAM) programme and will not be used for any other incompatible purposes. The data subjects will include: BAM participants, practitioners/counselors, school staff and parents/caregivers of BAM participants.

During this trial, data will be stored on secure, password-protected and encrypted network drives (hosted by BIT). Access to the data will be restricted to the relevant members of the project team involved in this evaluation.
Personal data will be subject to strict retention periods; the Lab will retain identifiable data only for the duration necessary to complete the analysis and reporting. Upon completion of the study, all data will be securely transferred via ONS to the YEF data archive (where YEF will act as the data controller for long-term archiving), and all locally held identifiable data will be securely and permanently deleted.

Furthermore, participant data rights will be actively upheld under the UK GDPR. Participants and their parents/caregivers will receive clear privacy notices detailing their right to be informed, to access their personal data, and to request rectification, erasure, or withdrawal from the study at any time prior to data anonymisation.

In case a Personal Data Breach occurs despite the mitigations in place, project team staff will deal with the security incident without undue delays. All personal data breaches (or suspected personal data breaches) will be reported to BIT's Data Protection Officer as soon as a project team member becomes aware of one (including if this is outside of office hours) by contacting the Data Protection Officer directly and by completing a Data Incident Notification Form. Staff will not attempt to investigate a personal data breach themselves but will take steps to contain the personal data breach as quickly as possible. Such steps might be taken prior to reporting the incident to the Data Protection Officer where this is reasonable and necessary to protect data subjects and mitigate the potential impact of the personal data breach. All anonymised data will be securely transferred via ONS to the YEF data archive at the end of the study and YEF will be a data controller for long-term archiving.

Stakeholders and interests

Table 9. Evaluation team.

Name	Role and responsibility	Institutional affiliation
Tom McBride	Project director. Responsible for ultimate project oversight and quality assurance.	The Ending Youth Violence Lab
Lilli Wagstaff	Project manager. Responsible for day to day project delivery, budget and timeline management.	The Ending Youth Violence Lab/The Behavioural Insights Team
Emma Forsyth	IPE delivery. Responsible for delivery of the IPE aspect of the pilot and efficacy trial.	The Ending Youth Violence Lab/The Behavioural Insights Team
Lisa Mueller	Researcher. Responsible for research set-up and delivery support.	The Behavioural Insights Team

Table 10. **BAM Delivery Team.**

Name	Role and responsibility	Institutional affiliation
Jane Caro	BAM Head of Service. Responsible for ultimate delivery oversight and quality assurance.	Islington Council
Ntale Eastmond	BAM Team Manager. Ensuring counsellors, school leaders, and parents have what they need to facilitate participation in the RCT.	Islington Council
Mariyam Farooq	Impact and Operations Lead. Serving as main/first point of contact for the evaluation team and supporting the rest of the programme team to complete all RCT related requirements.	Islington Council

Wendy Fine	Director of Research and Evaluation. Responsible for lessons learned and technical RCT expertise on BAM.	Youth Guidance
Jason Story	BAM Senior Replication Specialist. Responsible for training and coaching BAM Counsellors and supporting fidelity implementation.	Youth Guidance

Risks

Multiple potential risks of this study have been identified and relevant mitigations are being put in place. Please see the full risk register and mitigations attached. Key risks of the proposed research project are:

Table 11. Selection of potential risks.

Risk	Description and impact
Risks around self-reporting and sensitive outcomes	CYP or parents/carers may find reporting on sensitive outcomes offensive or distressing. It can also make them feel stigmatised and that others perceive them negatively. They may also fear that the information may be shared and there may be negative repercussions for them.
Contamination risk between trial arms	Contamination risk between trial arms Control group students, school staff, or the wider school environment may receive exposure to BAM content or mechanisms (e.g., if students share content or practitioners interact with control students), narrowing the measured difference between treatment and control groups (linking to RQ 5.2). Mitigated by monitoring contamination activities via the RAG rating framework in Appendix 5.
Participant adherence to the programme	Participants may not want to participate in all sessions, may have conflicts that make it hard to attend all sessions, or may move away from the local area which will affect

	our ability to detect impact. Participants may also wish to move between BAM circles, which would impact the dynamics of the groups. Decisions around circle formation are managed by counsellors.
Inconsistent intervention delivery	There are risks (in terms of interpreting study findings, and knowing what to potentially scale in future) around failing to deliver a consistent approach that is delivered across the different schools.
Low participation rates	Recruitment to trials can be challenging due to concerns around data collection, randomisation and the need for consents from young people and parents/carers.
Concerns from stakeholders, parents and CYP of randomisation	CYP, their parents / carers and referral partners may be wary of engaging with BAM due to the work involved with data collection, and risk of randomisation meaning that CYP might not be able to engage with the programme.
Sensitivity of police data linkage	Linking trial participants to police records is highly sensitive. CYP and their caregivers may be anxious about the implications of their records being accessed, fearing potential misuse, surveillance, or negative long-term impacts on future opportunities. This may lead to significant distrust in the evaluation process.
Participant perceptions of criminalisation	The use of police arrest data may lead participants or families to feel criminalised, or perceive the study as a form of institutional surveillance. This perception risks damaging the trusting relationship required for effective therapeutic engagement and may lead to participant withdrawal or refusal to consent.

Timeline

These are the timeframes of the key research phases:

Research phase/ activity	Timeframe
Pilot - project and evaluation set up and mobilisation	March 2026 - October 2026
Pilot - project and evaluation launch and delivery	September 2026 - April 2028
Anticipated progression decision	September 2027
Efficacy study - project and evaluation delivery	August 2027 - March 2030
Data management and analysis	December 2027 - June 2029
Final report (without 12-month police data)	April 2030
Final report with updated police data	November 2030

Please see attached gantt chart for a more detailed timeline.

Appendix 1 - Adaptation Approach

BAM London: 1-Year Curriculum Adaptation Principles and Process

Principles

- Ensure that curriculum meets the constraints of being delivered in 1-Year
- Adhere to the BAM Theory of Change
- Learn from previous BAM Evaluations to ensure core domains / practices of BAM are preserved in adaptation
- Maintain key aspects of program delivery, lesson framework, values
    - Check-ins, activities/portables, group missions
    - Core values – attempt to cover all six
    - Fun, safety, challenge
    - Maintain the **7 C's of Effective BAM Practice**: Curriculum Fidelity, Container of Safety, Core Values, Consistency, Code of Professionalism, Continuous Ruthless Self-examination/Journey to Self-Discovery, Community
    - Coaching
- Aim for deep understanding of BAM's core values, especially the first 3 core values (Integrity, Self-Determination, Positive Anger Expression) – keep the theoretical foundations and flow – maintain the intentionality
- Use data to inform critical curriculum elements

Process

- Broadly follow YG's Curriculum Review and Adaptation Framework *(See below)*
- Internal YG & MHF Committee meetings to map out adapted 1-year curriculum
    - Key lessons and activities
    - Lesson flow
    - After school/beyond school value internalisation demonstrations, enhancements & celebrations
    - Approach to continuing BAM connections post program year
- Get input from current BAM London youth participants
- Review YG BAM Curriculum tracking data to inform 1) which lessons/activities resonate most with students; 2) which lessons/activities are most consistently practiced by the best BAM counselors
- Review Participant Experience Survey qualitative data to identify core aspects of BAM that resonated most with youth and had the most impact:
    - What things they learned from BAM
    - What changes have they noticed in themselves since participating in BAM
    - What about BAM led to those changes

- Get feedback from Islington community
    - Compile new 1-year London BAMual

Timeline (January 2026 – May 2026)

Note: below is an outline of the initial time. Due to delays in the curriculum adaptation and delays in putting the BAM counsellors in place, activities related to feedback, publishing and training will require updates in the timeline. The full adaptation is expected to be completed in July 2026.

- **January 2026**
    - Review existing curriculum and theories of change, previous research to identify core program/curriculum elements to retain
        - BAMual, London BAMual, Advanced BAMual, BAM Justice Edition, BAM Virtual Curriculum
        - BAM Theory of Change, BAM London Theory of Change
        - BAM Qualitative Evaluation, BAM RCT Evaluations
        - Program Outcome Data (HSA, PES, Training & Coaching)
- **February 2026**
    - Obtain BAM participant input via survey or focus group
    - Propose Initial high-level Scope & Sequence for Adapted Curriculum
        - Assess proposed scope and sequence for alignment with adaptation process
    - Review and incorporate stakeholder feedback
    - Analyze Participant Experience Survey qualitative data and incorporate findings
    - Analyze BAM Curriculum tracking data and incorporate findings
- **March 2026**
    - Update Adapted Curriculum Scope & Sequence with all curricular elements to produce initial draft
    - Assess draft's alignment with YG's Curriculum Review and Adaptation framework
- **April 2026**
    - Co-Develop Curriculum Implementation Plan with London Delivery Team
    - Obtain & incorporate feedback from stakeholders on Adapted Curriculum Scope & Sequence
- **May 2026**
    - Publish & Share Adapted 1-Year Curriculum for BAM London (RCT Conditions)
- **June 2026 – June 2027**
    - Train and Implement 1-Year Curriculum
    - Monitor Implementation and Collect Feedback

  Revise and Update Adapted Curriculum for Use during full RCT

Program Structure

- Duration: One academic year
- Session Frequency: Weekly/Afterschool
- Session Length: 60–90 minutes
- Group Size: 8–12 participants, mixed tiers of need
- Facilitation: Trained BAM Counsellors
- Coaching: Regular coaching for BAM Counsellors provided by Curriculum Specialist/Replication Specialist

YG Curriculum Review Framework / Adaptation Process

- Prepare for review
    - Identify Review Team
- Assess each domain (4-Point Scale)
    - Objectives
        - Are learning goals clear, measurable, developmentally appropriate, and aligned with broader program mission?
    - Engaged Communities
        - Does curriculum include families, mentors, and community partners?
        - Are relationships reciprocal?
    - Competencies
        - Does it build real-world skills, mindsets and knowledge reflecting youth identity goals?
    - Cultural Conditions
        - Are materials relevant, affirming, stereotype-free, and responsive to youth realities?
    - Structure
        - Is it organized, sequential, and adaptable?
        - Does it fit time and program models?
    - Assessment of Competencies
        - Are assessments authentic, respectful, and youth-centered?
        - Are results used for improvement?
    - Quality Assurance
        - Is there a system for piloting, collecting feedback, and continuous iteration?
    - Youth Voice & Culture
        - Are youth co-creators, leaders, and visible in language, examples and choices?
    - Healing-Centered
        - Does it affirm humanity, promote belonging, and support collective healing?
    - Community Ownership
        - Are families and local stakeholders co-owners in shaping, delivering and refining curriculum?
- Engage youth and community input
    - Conduct youth and stakeholder questionnaires
- Document findings and plan revisions
- Implement revisions and training

Appendix 2: Theory of Change



Activities

Counsellors deliver a mix of structured and informal activities with young people. The core activity is the **BAM Circle**, a weekly peer support group of 8 to 12 members that follows a consistent format: a check-in (PIES), a **values-based activity** drawn from the BAManual curriculum, reflection, and application. Circles are supplemented by specialist sessions, including career workshops, martial arts, goal-setting, women-led talks, and guest visitors.

Beyond circles, young people take part in **monthly after-school sessions** that consolidate the month's value, **special activities** such as celebration trips, university visits, and an intra-BAM football league, **brief informal encounters**, and scheduled **1-2-1 support** with counsellors. Counsellors also provide extra support after a suspension, exclusion, or incident.

Delivery depends on: how consistently counsellors can run circles and specialist sessions, and how flexibly the curriculum is adapted to real-time incidents and individual student needs.



Mechanisms

These activities work by changing how young people relate to themselves and others. Circles build **self-regulation**: holding back from interrupting, letting calm spread through the group, and practising emotional expression safely. Counsellors also model alternative responses to anger, helping young people **build a relationship with a trusted adult**.

Roleplay and group missions prompt **perspective-taking and introspection**. Physically demanding activities and structured practice build a **sense of achievement**, and identifying with the group fosters a **sense of belonging**. Counsellors also teach young people questions to ask themselves in the moment of anger, **reducing automaticity** and creating a trained pause before reacting.

Moderating factors: the counsellor's own capability and consistency, whether the home/family environment reinforces these gains, and whether the physical space allows young people to disclose and engage safely.



Outcomes

In the short term, young people show **improved social skills**. As mechanisms take hold, this develops into medium-term gains in **self-esteem, self-control, and emotional regulation**.

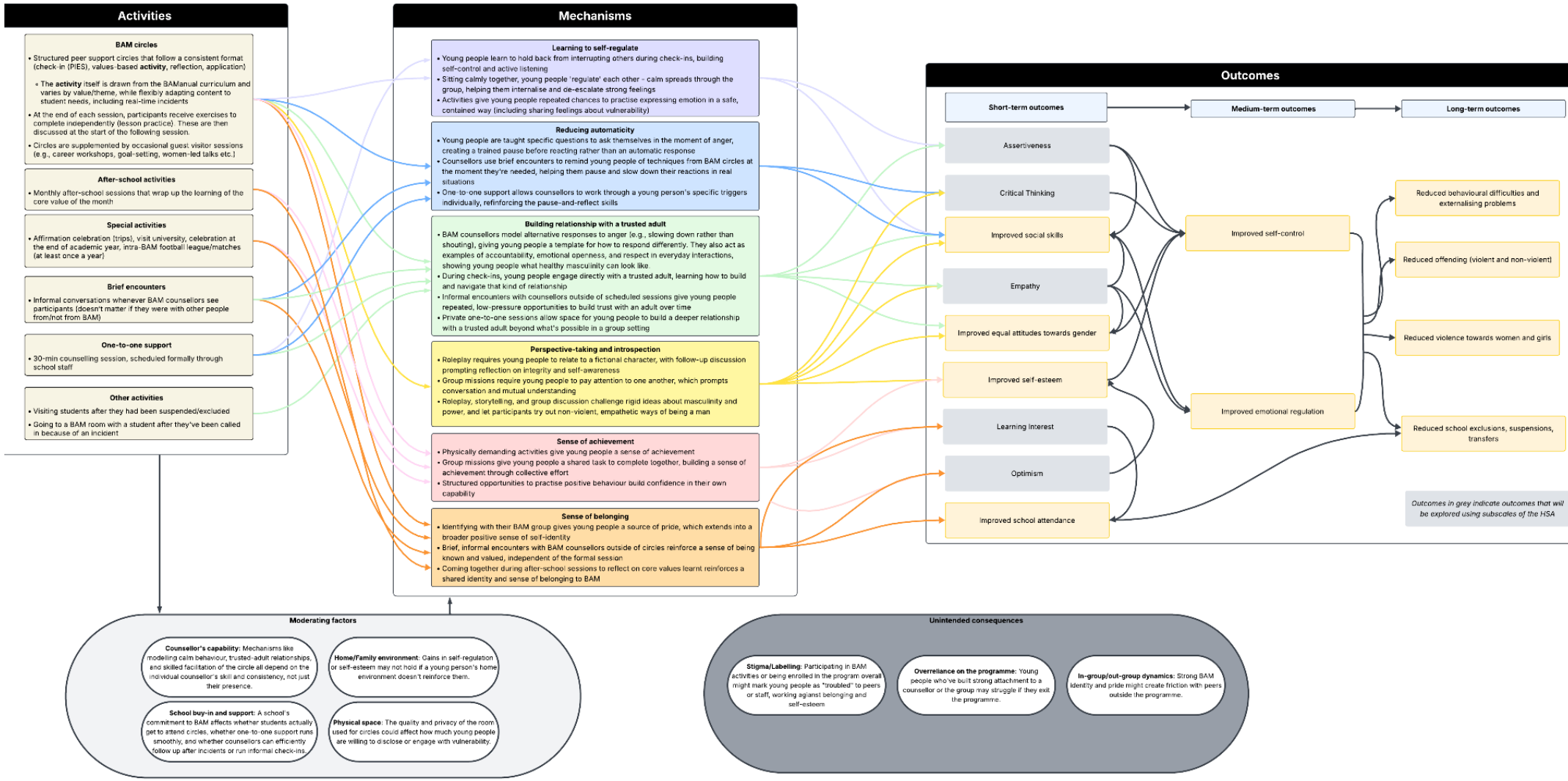
Over the longer term, these translate into **improved school attendance**, **reduced school exclusions, suspensions, and transfers**, **reduced behavioural difficulties and externalising problems**, and **reduced offending, both violent and non-violent**.



Unintended Consequences

The same mechanisms that help young people can also create risks. Being enrolled in BAM may lead to **stigma or labelling**, marking a young person as "troubled" to peers or staff. That works against the very belonging and self-esteem the programme aims to build.

Strong attachment to a counsellor or the group can create an **overreliance on the programme**, leaving young people to struggle once they exit. And a strong BAM identity can produce **in-group / out-group dynamics**, creating friction with peers outside the programme.


Activities
BAM circles
• Structured peer support circles that follow a consistent format (check-in (PIES), values-based activity, reflection, application)
◦ The activity itself is drawn from the BAManual curriculum and varies by value/theme, while flexibly adapting content to student needs, including real-time incidents
• At the end of each session, participants receive exercises to complete independently (lesson practice). These are then discussed at the start of the following session.
• Circles are supplemented by occasional guest visitor sessions (e.g., career workshops, goal-setting, women-led talks etc.)
After-school activities
• Monthly after-school sessions that wrap up the learning of the core value of the month
Special activities
• Affirmation celebration (trips), visit university, celebration at the end of academic year, intra-BAM football league/matches (at least once a year)
Brief encounters
• Informal conversations whenever BAM counsellors see participants (doesn't matter if they were with other people from/not from BAM)
One-to-one support
• 30-min counselling session, scheduled formally through school staff
Other activities
• Visiting students after they had been suspended/excluded
• Going to a BAM room with a student after they've been called in because of an incident
Mechanisms
Learning to self-regulate
• Young people learn to hold back from interrupting others during check-ins, building self-control and active listening
• Sitting calmly together, young people 'regulate' each other - calm spreads through the group, helping them internalise and de-escalate strong feelings
• Activities give young people repeated chances to practise expressing emotion in a safe, contained way (including sharing feelings about vulnerability)
Reducing automaticity
• Young people are taught specific questions to ask themselves in the moment of anger, creating a trained pause before reacting rather than an automatic response
• Counsellors use brief encounters to remind young people of techniques from BAM circles at the moment they're needed, helping them pause and slow down their reactions in real situations
• One-to-one support allows counsellors to work through a young person's specific triggers individually, reinforcing the pause-and-reflect skills
Building relationship with a trusted adult
• BAM counsellors model alternative responses to anger (e.g., slowing down rather than shouting), giving young people a template for how to respond differently. They also act as examples of accountability, emotional openness, and respect in everyday interactions, showing young people what healthy masculinity can look like.
• During check-ins, young people engage directly with a trusted adult, learning how to build and navigate that kind of relationship
• Informal encounters with counsellors outside of scheduled sessions give young people repeated, low-pressure opportunities to build trust with an adult over time
• Private one-to-one sessions allow space for young people to build a deeper relationship with a trusted adult beyond what's possible in a group setting
Perspective-taking and introspection
• Roleplay requires young people to relate to a fictional character, with follow-up discussion prompting reflection on integrity and self-awareness
• Group missions require young people to pay attention to one another, which prompts conversation and mutual understanding
• Roleplay, storytelling, and group discussion challenge rigid ideas about masculinity and power, and let participants try out non-violent, empathetic ways of being a man
Sense of achievement
• Physically demanding activities give young people a sense of achievement
• Group missions give young people a shared task to complete together, building a sense of achievement through collective effort
• Structured opportunities to practise positive behaviour build confidence in their own capability
Sense of belonging
• Identifying with their BAM group gives young people a source of pride, which extends into a broader positive sense of self-identity
• Brief, informal encounters with BAM counsellors outside of circles reinforce a sense of being known and valued, independent of the formal session
• Coming together during after-school sessions to reflect on core values learnt reinforces a shared identity and sense of belonging to BAM
Outcomes
Short-term outcomes
Medium-term outcomes
Long-term outcomes
Assertiveness
Critical Thinking
Improved social skills
Empathy
Improved equal attitudes towards gender
Improved self-esteem
Learning Interest
Optimism
Improved school attendance
Improved self-control
Improved emotional regulation
Reduced behavioural difficulties and externalising problems
Reduced offending (violent and non-violent)
Reduced violence towards women and girls
Reduced school exclusions, suspensions, transfers
Outcomes in grey indicate outcomes that will be explored using subscales of the HSA
Moderating factors
Counsellor's capability: Mechanisms like modelling calm behaviour, trusted-adult relationships, and skilled facilitation of the circle all depend on the individual counsellor's skill and consistency, not just their presence.
Home/Family environment: Gains in self-regulation or self-esteem may not hold if a young person's home environment doesn't reinforce them.
School buy-in and support: A school's commitment to BAM affects whether students actually get to attend circles, whether one-to-one support runs smoothly, and whether counsellors can efficiently follow up after incidents or run informal check-ins.
Physical space: The quality and privacy of the room used for circles could affect how much young people are willing to disclose or engage with vulnerability.
Unintended consequences
Stigma/Labelling: Participating in BAM activities or being enrolled in the program overall might mark young people as "troubled" to peers or staff, working against belonging and self-esteem
Overreliance on the programme: Young people who've built strong attachment to a counsellor or the group may struggle if they exit the programme.
In-group/out-group dynamics: Strong BAM identity and pride might create friction with peers outside the programme.

Appendix 3: Case Studies

Toby:

Toby is a Year 9 student. Toby is excellent at art, enjoys science experiments, and is generally known by staff as polite.Toby has recently been flagged by his form tutor as needing general support to improve his wellbeing. He does not have anything on his behaviour record but over the past term, his grades have fallen. Toby's parents think he is feeling low in mood because of recent friendship problems.

Why he is included: Toby has been identified by school staff as needing general support for his mental health because of recent changes in behaviour.

Marcus:

Marcus is a Year 8 student. He is creative, loves working with computers, and is known for being very kind. Marcus's older brother has recently been arrested, which has brought a lot of stress into his home. Because of this stress, Marcus has been flagged by the pastoral lead as he appears to be more anxious and has started to argue more with his teachers.
Why he is included: The justice system involvement of a close family member and staff's concerns of his wellbeing and behaviour.

Tyler:

Tyler is a Year 8 student. He is a talented graphic artist, loves running, and is highly independent. Tyler processes things a bit differently and can find loud, busy classrooms very overwhelming. He was recently suspended after his frustration built up during a lesson, where he walked out of class and refused to come back. Tyler tends to cope with his stress using substances with older boys on his street, which raised concerns from school.

Why he is included: Tyler has a history of suspension and has known substance use, with a risk of being taken advantage of by those in his area.

Jordan:

Jordan is a Year 10 student. He has a big friendship group, loves fixing bikes, and is passionate about music. Jordan was recently arrested after a fight in his neighborhood and received a formal caution from the police. Because of this incident, he is currently receiving support from the local "Youth Offending Team" but not in any therapy. The stress of this situation has made him feel isolated and defensive at school, leading to a drop in his attendance and grades.

Why he is included: Jordan is currently involved with the local youth offending team and has received a police caution.

Appendix 4: Referral / Recruitment Rate Table

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Months:	April - June 26	July - Sept 26	Oct-Dec 26	Jan - March 27	April - June 27	July - Sept 27	Oct-Dec 27	Jan - March 28	April - June 28	July - Sept 28	Oct-Dec 28	Jan - March 29	April - June 29	July - Sept 29
Target number of children and young people **referred into the project**	0	270	0	0	0	540	0	0	0	540	0	0	0	0
Target number of children and young people **recruited to the project and evaluation**	0	216	0	0	0	432	0	0	0	432	0	0	0	0
Expected number of children and young people who **withdraw/drop out of the project**	0	27	9	9	9	54	18	18	18	54	18	18	18	18
Expected number of children and young people **who drop out of the evaluation**	0	0	0	0	0	21.6	0	0	0	43.2	0	0	0	43.2

BAM Contamination Risk: RAG Rating

This document provides a RAG (Red/Amber/Green) rating of contamination risk across BAM programme activities. Contamination refers to control group students, staff, or the wider school environment receiving meaningful exposure to BAM's content or mechanisms, which would bias evaluation findings by narrowing the measured difference between treatment and control groups.

Risk is considered across the full evaluation context: programme delivery, school environment, staff and curriculum, data collection, and informal or peer-driven exposure.

Rating Definitions

Green	**Low or negligible contamination risk** *Awareness of BAM's existence or presence without receiving any of its content* These activities are safe for counsellors to continue to deliver
Amber	**Moderate or indirect contamination risk** *Control group students receive partial or indirect exposure to BAM* These activities will require careful thought, consideration and guidance from either YG or EYV Lab before being delivered
Red	**High contamination risk** *Control group students receive direct exposure to BAM* These activities should never be delivered

RAG Ratings

1. Programme delivery

Green	• General staff or school awareness that BAM operates in the school, without adopting its content
Amber	• Special activities with a mixed audience (e.g. an employer workshop open to the whole year group rather than BAM only)
Red	• BAM activities held in shared or semi-public spaces where

	control group students can regularly overhear content • Control group students included in BAM circles • Control group students attending BAM special activities alongside participants (trips, football league, celebrations) • Control group students receiving 1-2-1 counselling support from a BAM counsellor • A BAM counsellor providing BAM-style support to a control group student after suspension or exclusion

2. School environment

Green	• The physical BAM room being known or visible in the school, without control group students entering it or receiving support there
Amber	• Teachers or other school staff picking up BAM language or techniques informally and reusing them with control group students • BAM counsellors delivering training or talks to non-BAM staff, who then carry that content into their own interactions with control group students • BAM students informally sharing core values, anchor questions, or check-in habits with control group friends • Control group students socialising closely with BAM students such that they absorb BAM language, norms, or techniques over time
Red	• The school adopting BAM's core values language or anchor questions as official whole-school behavioural policy • A PSHE course or similar for all students that directly teaches BAM's core values, anchor questions, or curriculum content

Parking lot

Please use this space to list any other relevant activities that should be included in the RAG rating. The EYV Lab will regularly check this space and RAG rate activities, moving them into one of the tables above.

Interactions with control group pupils during break time or lunchtime. For example, supervising football at lunchtime.

Saying hello to control pupils in the corridors/hallway (not any deeper conversation, just a hello, how are you?)

YOUTH ENDOWMENT FUND

Impetus

Home Office

youthendowmentfund.org.uk

hello@youthendowmentfund.org.uk

@YouthEndowFund

The Youth Endowment Fund Charitable Trust

Registered Charity Number: 1185413