

EVALUATION STRATEGY

Another Chance Programme Formative Evaluation Strategy v4.0

University of Hull, UCL – Security & Crime
Science and Abertay University

Paul McFarlane, Will Graham, Iain Brennan
Alex Sutherland

October 2023



Document Version Control and Review

Version No:	4.0
Effective Date:	25th October 2023

Authorship and Approval		
Name and Designation	Author/Reviewer/ Approval	Date
Evaluation Team	v1.0 Author(s)	31st October 2022
	v1.0 Reviewer	
	v1.0 Approval	

Document Revision History		
Version No	Date	Summary of Revisions
1.0	31st October 2022	Initial creation/New document
2.0	7th December 2022	Creation of v2.0 addressing v1.0 feedback comments
3.0	4th September 2023	Creation of v3.0 with project updates
4.0	25th October 2023	Creation of v4.0 from YEF comments on v3.0

Programme Formative Evaluation Strategy v4.0

1.0 Introduction

This final revision to the programme's formative evaluation strategy complements the published summative evaluation protocol, reflects the progress made in the preparation and early implementation phases of programme delivery at each site and deals with Youth Endowment Fund (YEF) comments on v3.0. It represents a better understanding of the theoretical dimensions, mechanisms and limitations of both the intervention and the evaluation.

To date, the programme's primary delivery activities have been focused on developing the treatment condition at the five sites located in Coventry, Leicester, Manchester, Nottingham, and Wolverhampton. In each case, the intervention has been designed in alignment with the nine criteria outlined in the Youth Endowment Fund Focused Deterrence Framework. Participants have been randomly assigned to either the treatment or control (business-as-usual) conditions. This assignment was based on pre-established inclusion criteria relating to participants' histories of involvement in violence.

Aligned to realist principles, the framework for the evaluation of the treatment condition has now been developed and is structured around three interconnected configurations: (a) the 'targeted enforcement' mechanism; (b) the 'individualised support' mechanism; and (c) the 'community validation' mechanism (see Section 1.7).

1.1 Background

The Agency Collaboration Fund 1 (ACF1) is a multi-site, multi-partner programme commissioned to evaluate how the intervention, focused deterrence, affects violence in high-risk population groups. This document describes the programme formative evaluation strategy, i.e., the high-level evaluation framework that will be relied on to explain how the programme was developed and implemented in various local contexts.. The evaluation strategy is based around two key questions: (i) what are the important components that contribute to the successful implementation of a focused deterrence framework (e.g., ways of operating, partnership models, community engagement) to reduce youth violence? and (ii) how might a focused deterrence framework be utilised most effectively in various socioeconomic and local contexts to reduce youth violence?

In terms of the wider context, we recognise that disparities are present across various sectors in the UK, with individuals from Black, Asian, and Minority Ethnic (BAME) backgrounds experiencing overrepresentation, particularly in the criminal justice system. Studies indicate that Black individuals are more likely to be victims of violent crime and to be stopped, searched, arrested, and imprisoned than their White counterparts. This pattern of overrepresentation is often associated with systemic biases and socioeconomic conditions that have historically affected minority communities. Similarly, in education, students from particular ethnic minorities and people from lower socioeconomic backgrounds often face disparities in the provision

of education. This includes lower levels of attainment, higher rates of exclusion, and less access to higher education opportunities. Language challenges, cultural differences, and a lack of focused support for these children are all factors that contribute to these variations. In employment, inequities are also visible, with BAME people experiencing higher unemployment rates and being underrepresented in senior positions. Even when working, they frequently face a salary disparity compared to White coworkers.

Another source of concern is health inequity, with minority groups and people from disadvantaged areas facing poorer health outcomes, reduced life expectancy and limited access to healthcare services. The COVID-19 pandemic highlighted these inequities, with BAME populations facing greater infection and fatality rates due to factors such as frontline employment, multigenerational households and pre-existing health problems. To address them, the government, institutions and communities must work together to establish more inclusive policies and interventions that are customised for all individuals, regardless of their background.

Given this, the programme's complexity poses a risk for fully knowing how the intervention might work, for whom, and in what context. Although the literature on multi-partner violence prevention initiatives in the UK is limited, the wider literature on complex interventions demonstrates the importance of acknowledging that an intervention can work through multiple latent mechanisms, each of which can be affected by changing contexts and local conditions in which the intervention is delivered. For a community violence prevention intervention, these contextual factors include the effect of interactions between and within local systems (e.g., local political support, local partnership structures, governance and power dynamics, availability of appropriate levels of funding for stakeholders), social situations (e.g., rates of community violence, disproportionality, levels of investment in services), norms (e.g., community trust in statutory services, especially law enforcement), values (e.g., street codes, commitment to family) and relationships (e.g., history of local engagement with services, local access to resources, population demographics). All of these characteristics have the potential to randomly interact, affecting the way in which a planned intervention is implemented at the sites and being experienced by population groups and communities that might be disproportionately exposed to violence.

Although impact evaluation via a randomised or quasi-experimental study design can give valid estimates of the effect of an intervention on an outcome, without a concomitant understanding of local context and implementation, the potential for wider scaling of an intervention is limited. To address these challenges, we have selected a realist randomised controlled trial evaluation strategy (Bonell et al., 2012) that combines the nuance of realist evaluation with the robust inferences afforded by a randomised controlled trial methodology. In brief, realist evaluation is a theory-based evaluation method that emphasises understanding how an intervention causes change and, as far as possible, describing how that causal mechanism varies across contexts and populations. Realists may suggest that an estimation-focused method such as a randomised controlled trial is epistemologically incompatible with a theory-based method such as realist evaluation. However, recent advances in evaluation research have proposed setting aside these disputes in favour of a pragmatic solution where a

realist randomised controlled trial offers the greatest insight into how and if a complex intervention works. We will use the tools of realist evaluation for both the programme-level and site-level formative evaluations.

1.2 Overview of focused deterrence

Focused deterrence, often referred to as ‘pulling-levers’ was developed as a problem-oriented policing project to deal with serious gang violence in Boston in the mid-1990s by a team of academics. Focused deterrence interventions target “very specific behaviours by a relatively small number of chronic offenders who are highly vulnerable to criminal justice sanctions” and “directly confront offenders and inform them that continued offending will not be tolerated and how the system will respond to violations of these new behaviour standards”.

The concept of general deterrence is central to our understanding of how risks are perceived by offenders by deploying traditional and non-traditional law enforcement tools to directly communicate incentives and disincentives to targeted offenders (Kennedy 1997). These incentives can take a variety of forms, but most often involve local support to disengage from violent groups and to desist from violence (Braga et al., 2019). Support can be provided by a range of individuals as well as organisations in the statutory, voluntary and private sectors. Disincentives are usually provided by law enforcement agencies but can involve a range of other ‘levers’ available to statutory services, such as the legal denial or rescinding of existing support. As such, focused deterrence can have a multi-agency component that draws together a wide selection of resources that may support disengagement and desistance.

The intervention framework employs both incentives and disincentives, reflecting its multifaceted nature. Multiple organisations will deliver various components of the intervention, which can work together or operate independently. This multi-agency, multi-component approach adds a high degree of complexity to the intervention. The systems activated to change individual violent behaviour in the community are numerous and were not initially designed to work together. This context, with its interdependencies among social, cultural, economic, and political factors, is characteristic of a complex intervention. Understanding this complexity is crucial for our evaluation strategy, and it is integrated throughout. Our current grasp of the intervention's complexity is detailed in Section 3.3 and is also theorised in the programme-level theory of change. We anticipate that the data gathered from sites will help to improve and enrich our understanding of the complexity of the intervention as we progress through the early implementation phase.

1.3 Evidence base and theoretical limitations

Our current understanding of the evidence base on the likely effects of focused deterrence is that it is based on weak study designs. Braga et al.'s 2019 systematic review concluded that while focused deterrence may be effective, the evidence base needs to be more rigorous and the causal contribution of each component of the intervention needs to be better understood. However, none of the studies included in the review were randomised controlled trials, pre-registered, or pre-specified and most had too small sample sizes to make reliable population inferences. Furthermore, we

take the view that the studies varied in terms of the outcomes used, levels of aggregation, and inferences made. The treatment units in prior research were typically groups ranging in size and structure, and the outcome measures used varied from narrowly to broadly defined violent crimes. The comparison groups also varied depending on the outcome being analysed. The studies are highly susceptible to biases, including false positive results and incorrect attribution of regression to the mean as treatment effects (see protocol for more detail). A significant publication bias also exists, which, when adjusted, diminishes the effects of these quasi-experimental studies. As a result, claims of a strong evidence base for the effectiveness of focused deterrence are overstated. As such, one of the goals of the realist design of this programme formative evaluation strategy is to develop a more nuanced understanding of the contribution of each of the three components to any observed effects.

Table 1: Phase names for the programme evaluation

Phase	Name	Date
Phase 1	Coalignment	April 2022-June 2022
Phase 2	Preparation	July 2022-May 2023
Phase 3	Delivery Early implementation - Pilot	June 2023-November 2023
Phase 4	Delivery Full implementation	December 2024-December 2025
Phase 5	Follow-up and write-up evaluation	December 2025-July 2027

1.4 Aims of programme formative evaluation strategy

As set out in previous versions, the aims of the programme formative evaluation strategy are aligned with the three principles of YEF evaluation, which are: (i) developing a better understanding of the evidence for implementing focused deterrence in a UK context; (ii) producing replicable lessons for future interventions; and (iii) optimising the fund’s value for young people at risk of violence. In accordance

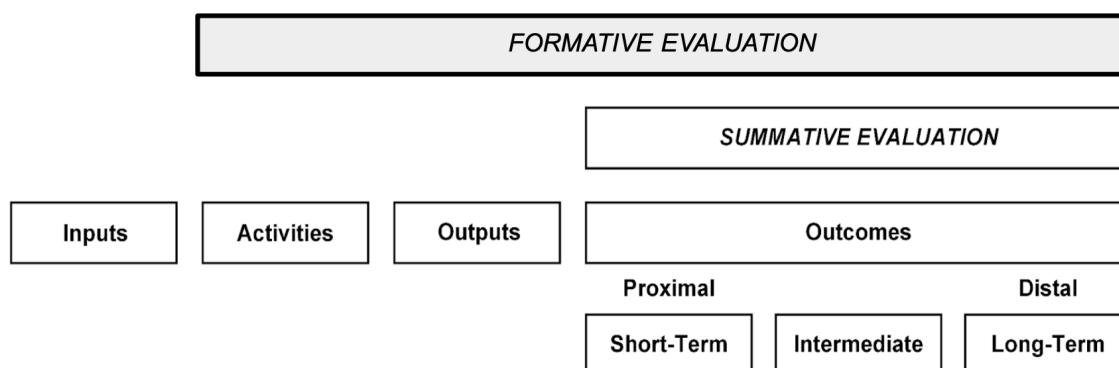
with these principles, this strategy has, to date, generated the following programme-level output:

- Co-creation of a (i) programme-level theory of change and (ii) site-specific theories of change that reflect pre-implementation expectations about the intervention (shared with YEF);
- A narrative description of implementation plans of the intervention across multiple sites, including information about detailing the fidelity of the implementation of the YEF framework (included within protocol);
- Early Identification of issues and lessons learned related to theories of change and practical implementation of the intervention;
- Providing YEF with clear and timely reporting on progress with sites to troubleshoot problems early and produce a clear account of site-level implementation (system health checks); and
- Developing a rubric to detect early progress and decide whether proximal outcomes are trending in the desired direction (see table of proximal outcome measures in Appendix B).

1.5 Scope and operationalisation of the evaluation strategy

Figure 1 sets out how the scope of the programme formative evaluation strategy relates to the programme theory of change inputs, activities, outputs and proximal (short-term) outcomes. Typically, a formative evaluation will consider the effect of the intervention on proximal outcomes. However, at the request of YEF on review of version 1, process data will be collected related to the effect on intermediate and long-term outcomes.

Figure 1: Scope of formative evaluation



The scope of the strategy will be operationalised by nine realist formative evaluation questions (FEQ). The approach, data source and methods to answer each of these questions are set out in Section 2 below.

The programme level formative evaluation questions are:

FEQ 1: To what extent were the three components of the intervention, as required by the YEF framework, received by the treatment population groups?

FEQ 2: How did inputs contribute to the intervention functioning?

FEQ 3: Who did the intervention work for and how?

FEQ 4: How did local context affect intervention delivery?

FEQ 5: To what extent was the intervention delivered as intended (Fidelity)?

FEQ 6: How did complexity affect intervention delivery?

FEQ 7: How did proximal outcomes change?

FEQ 8: Why did proximal outcomes change?

FEQ 9: What was learned from how the intervention was delivered?

1.6 Target population

The population groups targeted by the ACF1 YEF FD intervention are 'those associated with groups involved in serious youth violence.' In this case, "youth" refers to those individuals between the ages of 10 and 24, regardless of sex, gender, ethnicity, or social background. However, the study now includes individuals ranging from 14 to 40 years old, extending beyond the YEF's defined age limit for youth. This broader age range has been selected because focused deterrence interventions typically target adults, and the intervention's effects may differ across age groups. To address both components of the YEF remit, the decision was made during the preparation phase to include both children and adults in the study, and as such, the formative evaluation will examine differences in intervention effects and experiences between these two groups.

1.7 Programme theory and evaluation approach

The programme formative evaluation strategy adopts a realist approach and will assess the components of the intervention and measure them against the defined proximal outcomes across varying local contexts. Central to this strategy is the development of a pre-defined programme theory of change (P-ToC). The P-ToC, was co-created by evaluators and implementers during Phase 2 and incorporates the requirements related to nine core elements of the YEF FD framework, making allowances for local complexity and context. Based on the research that has already been done on focused deterrence, the P-ToC explains what we know at this stage about the theoretical processes that are supposed to make the change that the intervention wants to happen.

A key component of the realist nature of this evaluation is the high-level context-mechanism-outcome (C-M-O) configurations that have been co-created prior to the early implementation phase. These C-M-O configurations provide the analytical basis for the formative evaluation strategy. In this instance, the mechanism component is not a variable but is used to explain anticipated participant reasoning and responses to the deployment of intervention resources and activities in different contexts. In other words, mechanisms provide credible assertions regarding generative causal links between intervention inputs, activities, outputs, and likely observable effects (e.g., outcomes and impact).

As recommended by an external expert (Prof Chris Bonnell, London School of Hygiene and Tropical Medicine, UK), the evaluation will use a limited number of C-M-O configurations centred on generally accepted 'big' ideas that have been previously used in the literature to plausibly explain three ways in which the focused deterrence intervention might activate new mechanisms to inhibit violent crime-generating processes. This includes targeted enforcement, the provision of individualised support, and community validation of the intervention.

As shown in Table 2, these propositions have been developed into three high-level C-M-O configurations: (a) the targeted enforcement mechanism; (b) the individualised support mechanism; and (c) the community validation mechanism. These configurations will be used to capture the complexity of the intervention and develop a set of downstream realist causal propositions that can be tested in local 'observable' contexts. These C-M-O configurations will thus be used to provide the framework to guide data collection, according to realist principles.

Table 2: High-level C-M-O configurations (v2.0)

Context +	Mechanism =	Proximal outcome(s)
Heterogeneity in local resource availability and deterrence activities	[Targeted enforcement] An increase in legitimate targeted deterrence activities affects perceptions of certainty of arrest and punishment	Normative and instrumental non-violent modifications and engagement with individualised support
Variance in resource allocation, coordination, engagement, collaborations and spectrum of available support services	[Individualised support] Increase in availability of improved individualised support packages, providing pathways to desistance and increasing perceptions of benefits	Sustained engagement with pathways to desistance and reorientation towards legal and social norms
Decrease in levels of community confidence in local policing and statutory and non-statutory support services	[Community validation] Social amplification of community moral voice, characterised by peer and familial influences, informal social controls, collective efficacy, and communicating shared values and beliefs, affects normative compliance and behaviour modification	Increase in legitimacy and support for intervention and enforcement activities to achieve normative compliance

Table 3 shows how the YEF core elements have been aligned to the three high-level conceptual mechanisms: targeted enforcement, individualised support, and community validation.

Table 3: Core components of YEF FD framework

YEF Core Component	Mechanism(s) (to deliver essential component)
Focused deterrence must address serious youth violence where there is a clear group dynamic	Individualised support; community validation; targeted enforcement
Focused deterrence must be targeted to those associated with groups involved in serious youth violence	Individualised support; community validation; targeted enforcement
Focused deterrence must include an inter-agency working group made up of law enforcement, community and support services representatives	Individualised support; community validation; targeted enforcement
Focused deterrence must include high-quality, structured intelligence gathering and analysis	Individualised support; community validation; targeted enforcement
Focused deterrence must include a special enforcement operation that is swift, certain and includes group consequences	Targeted enforcement
Focused deterrence must include direct and frequent communication with target groups	Individualised support; community validation; targeted enforcement
Police partner: to deliver open, transparent and respectful policing and communicate swift, certain, and credible consequences of violence	Individualised support; community validation; targeted enforcement
Support services partner: to coordinate and mobilise a breadth of services with fast response times	Individualised support
Community partner/s: to express norms, expectations, and aspirations, support violence prevention and increase engagement in support services.	Individualised support; community validation

1.8 Primary, secondary and proximal outcomes

The programme evaluation is now focused on collecting data related to the theorised causal linkages between implementation activities, outputs and outcomes set out in the programme-level theory of change. In terms of evaluating overall programme impact, the primary outcome measure is the number of offences of violence against the person attributed to an individual within one year of randomisation. Secondary outcome measures include: (i) the number of days between randomisation and a recorded offence of violence against a person with a PNC disposal outcome relevant to the evaluation; and (ii) when there is a co-offender, the number of 'any' recorded offences attributed to an individual with a PNC disposal outcome relevant to the evaluation.

At the end of the pseudo-pilot and at six-monthly intervals during full implementation phases, the evaluation team will collect data related to proximal outcome measures (see Appendix B) to observe whether the site interventions and likely effects are trending in the desired direction. Proximal outcome data will be collected through interviews with participants and project team members and from administrative data. As shown in Appendix B, the formative evaluation strategy will use a common set of proximal outcomes to facilitate site-to-site comparisons and maximise early learning opportunities.

1.9 Inclusivity, equity and cultural sensitivity

The programme formative evaluation strategy acknowledges that the implementation of the intervention in behavioural and social settings that typically have higher rates of violence may result in stigmatisation or disproportionate targeting of individuals, groups, or communities. Consequently, the evaluation will collect preliminary data on evidence of inclusivity and the risk of racial disparity during early and full implementation (Phases 2 and 3).

This information will be used to examine proportionality and racial equity across the life cycle of the programme including (i) eligibility criteria, (ii) selection and referral of individuals, (iii) assessment outcomes, (iv) engagement activities and outcomes, (v) referral activity to support and (vi) to deterrence pathways, (vii) engagement with support services, (viii) perceived suitability of those services, (ix) deterrence activities, (x) perceived fairness of those deterrence activities and (xi) graduation and exiting of the programme.

Data for these will be drawn from a combination of process data and cohort interviews. Denominators will be derived from a combination of census and open-source police data. In several cases, the denominator for these judgements is itself likely to be racially biased. In such cases where adjustment for such biases is not possible, such as baseline patterns in rates of criminal charge across ethnic groups, such biases will be noted and carried forward to inform subsequent judgements of proportionality and bias in the pathway and in the use of statistical models of the trial outcomes.

The evaluation team has engaged a researcher with significant experience of researching with young black men, including those involved with the criminal justice system to lead team workshops on researching with this population and to support the development of appropriate research questions, methodologies and analytic techniques. This researcher's support will be fundamental when analysing the qualitative data obtained through interviews with the intervention cohort. While the research team cannot adopt a sufficiently nuanced racial lens through which to interpret the data, they will have a partner who can.

Through consultation with LERGs and race equity advisers, research questions will be critically reviewed for the presence of implicit bias or inappropriate assumptions. All research materials will be assessed in terms of their appropriateness for the sampled population. In particular, sampling methods and materials for interviews with a sample of the target population will be subject to ongoing scrutiny.

The evaluation will proactively try to include participants from all groups within local communities while acknowledging the difficulties associated with realist evaluations and research more generally. For instance, the evaluation team has been engaging with members of local communities by attending Independent Advisory Group (IAG) meetings in order to seek their support and guidance and gain a better understanding of how to identify and remove obstacles to engaging with hard-to-reach groups (e.g., legitimacy, trust, cultural sensitivities, language). To support this criterion, it is intended that IAGs appoint members of the local community as "pathfinders" to work together with the evaluation team.

In addition to engagement with IAGs, local evaluation reference groups (LERG) and expert advisors have been consulted on all evaluation questions, measures and interview schedules, with an emphasis on inclusiveness and equity. The evaluation is guided by the principles of trustworthiness, neutrality, voice, and respect, and all evaluation materials and processes will be prepared and reviewed with these values in mind. Evaluators will work with sites to help mitigate the occurrence of targeting or affecting any group disproportionately. Defining differential engagement patterns is rather simple; however, identifying targeting patterns is more difficult and will require access to more data, which will be discussed in ongoing interviews with delivery teams, stakeholders, and participants.

1.10 Outputs and deliverables

All outputs will be delivered in accordance with the terms of the YEF Project Evaluation Agreement dated 25th July 2023.

Table 4: Phases of programme evaluation strategy

Phase	Evaluator Activity	Output ¹
1. Baseline starting conditions 2. [completed]	<ul style="list-style-type: none"> ● Review baseline starting conditions and local context by conducting site-level interviews with stakeholders ● Draft programme theorised description of the intervention ● Assess local descriptions of intervention, causal assumptions, theories of change and logic models against YEF FD framework to identify gaps in theory, logic and assumptions ● Complete benchmarking of inputs, outputs and proximal outcomes to generate baseline indicators and make assessment of potential issues related to race equity in evaluation methods and data collection (see Section 2.5) ● Critically evaluate and address potential racial biases in research questions and measurements 	1;2
3. Early implementation 4. [on-going]	<ul style="list-style-type: none"> ● Review early implementation phase ● Assess whether proximal outcomes are trending in desired direction ● Make recommendations for programme functioning and quick improvements to delivery of intervention ● Review programme theory of change and C-M-O configurations ● Recommend to YEF suitability for progression to full implementation ● Assess race-equitable delivery of early implementation and feed back observations to YEF and sites 	3; 4; 5; 6
5. Full implementation 6. [awaits]	<ul style="list-style-type: none"> ● Collect longitudinal data using range of corroborative methods ● Monitor race-equitable delivery of intervention and incorporate assessment in feedback to sites and YEF 	7; 8
7. Progression to summative evaluation 8. [awaits]	<ul style="list-style-type: none"> ● Complete formative evaluation of programme and each project site ● Complete pre-post-programme theory of change ● Recommend to YEF progression to summative evaluation 	9

¹ See Section 1.10 for programme evaluation strategy outputs

2.0 Formative evaluation design

The formative evaluation will employ a mixed-methods design combining qualitative and quantitative methods to answer the formative evaluation questions (see Section 1.5). This design has been approved by the University of Hull's Faculty of Arts, Cultures and Education Ethics Committee. The rationale for selecting this design is that early qualitative data can and has now been used iteratively to discover and theorise important context-mechanism-outcome configurations. These configurations are used to develop hypotheses about what might work, for whom, and in what circumstances. Data will be collected and analysed on these configurations to answer the evaluation questions. Additionally, qualitative methodologies will be employed to record developing changes in implementation, intervention experiences, and unanticipated or complex causal pathways. Equally, the early identification of implementation and participation barriers using quantitative administrative data is also a crucial aspect of the approach for evaluating the programme intervention. For instance, if routine monitoring data indicates that the intervention is not reaching targeted population groups or that relatively few members from minority ethnic groups are not participating, interviews, focus groups, or self-completion questionnaires can be undertaken with stakeholders and potentially participants to identify and implement measures to mitigate process barriers and moderate these effects of the intervention.

2.1 Data collection methods

The evaluation team researchers will use a range of methods to collect longitudinal, cross-sectional, and pre-post data from multiple corroborative data sources, including delivery stakeholders, participants, project documents, administrative records, etc.,. We anticipate that, for practical reasons, data collection will be weighted towards the use of semi-structured interviews and observations. Below are the methods we are likely to use during the formative evaluation (we do not offer descriptions and definitions of these methods):

- Semi-structured interviews
- Ethnographic observation
- Routine performance monitoring using administrative data
- Survey questionnaires

Table 5 provides an overview of the implementation and process evaluation methods and provides a comprehensive guide for data collection from each of the groups involved in the intervention. It outlines distinct approaches for adult and child participants, project leads, the delivery team, and other stakeholders, both statutory and voluntary. The table specifies the type and number of participants (though numbers are yet to be determined), the methods of data collection (interviews or observations), and the settings in which these data will be collected. It also sets out the frequency of data collection, which is either cross-sectional or longitudinal, and the analytical methods to be employed, primarily thematic analysis and descriptive statistics.

Importantly, Table 5 aligns the formative evaluation questions (FEQs) relevant to each sample group, aimed at exploring critical programme dimensions such as the intervention's effectiveness, the role of local context, the contribution of various inputs and the complexities affecting delivery. In each sample group, researchers will use customised sets of questions during interviews or observations, designed specifically to align with the overarching FEQs. These question sets, containing a list of exploratory themes, are derived from the programme's theory of change and are further nuanced by a series of sub-evaluation questions (SEQs), which are set out in Appendix A. The SEQs serve a complementary role by offering detailed, context-specific insights to bolster the broader themes explored through the FEQs, increasing the likelihood of generating a comprehensive understanding of the theorised C-M-O configurations and the intervention's impact.

2.2 Data sources

We are aware of the well-known shortcomings and biases related to police-generated data. The evaluation will therefore gather administrative and routine performance data from various sources. These sources include police records, crime reports, meeting minutes, site management records, school exclusions, community impact assessments, community surveys, health records, demographics and more. Each data source will comply with site-level data sharing agreements and data protection impact assessments.

2.3 Analysis

The qualitative and quantitative data will be analysed to develop a common understanding of how, why and for whom did the intervention work for in varied local contexts. The flexibility and depth of the longitudinal qualitative data that will be collected is a fundamental strength of the programme formative evaluation design, allowing evaluators to investigate complex processes and unexpected outcomes from a realist perspective. It will be an iterative process to collect and thematically analyse these qualitative data against a defined conceptual framework. Theoretically, this means that emerging inductive topics can be explored in greater detail throughout data collection during the project lifecycle. Where practicable, the demographic profile of researchers collecting data will be representative of the participant population group.

Specifically, qualitative data will be entered into the NVivo software tool. The qualitative data will be thematically analysed. A set of a priori codes will be utilised to identify both emerging and recurring themes and concerns. This will be an iterative process in which these themes will also be utilised to analyse and compare data throughout the project's lifecycle, thereby establishing strong links between these data and future theories or hypotheses. These data will also be combined with other evaluation components. In this case, the objective is to leverage quantitative data collected on process components such as fidelity, hypothesised configurations of contextual influences, mechanisms, and outcome measures.

To ensure the validity of the formative evaluation, several steps will be taken to minimise researcher bias. First, all data will be anonymised and analysed against a

pre-defined conceptual framework, ensuring a consistent approach and limiting personal interpretations. Second, an iterative approach will be employed, allowing for the refinement of insights and ongoing checks of emerging themes. Third, every set of data will be analysed by two independent researchers to provide multiple perspectives. Fourth, all researchers will have undergone training on race equity and bias awareness with a subject expert and will engage in reflexive and equitable practices to mitigate potential biases. Fifth, if needed, an internal peer review process will be put in place to question and confirm interpretations. This will make sure that we have a complete and fair picture of how the intervention worked.

In regard to the summative component, qualitative process analysis data may serve predictive or post-hoc explanatory purposes. In other words, if these data are analysed prior to the summative analysis, they could provide prospective insights into why evaluators may anticipate favourable or negative overall intervention effects. For example, longitudinal qualitative data collected relating to proximal outcomes should, in advance, indicate whether there are likely to be observable effects on longer-term outcomes. Qualitative data may also contribute to the creation of hypotheses regarding the emergence of variety in outcomes; for instance, if particular groups of young people or communities appear to have responded to the intervention or its components better than others. Similarly, these data can be used as potential indicators for negative disparities latent within the intervention. Community stakeholders will be provided with the opportunity to be involved in interpreting the data in their local contexts. Capturing these data is vital to the programme's fidelity.

Table 5 : Implementation and process evaluation methods overview ²

Sample group	Type	Collection method	Location	Frequency	Analytical method	Formative evaluation question
> 17-year-old participants	Treatment group	Interview	Home, community or delivery team setting	Cross-sectional / longitudinal	Thematic analysis	FEQ 3: Who did the intervention work for and how? FEQ 4: How did local context affect intervention delivery?
<18 participants	Treatment group	Interview	Home, community or delivery team setting	Cross-sectional / longitudinal	Thematic analysis	FEQ 3: Who did the intervention work for and how? FEQ 4: How did local context affect intervention delivery?
>17 participants	Treatment group	Observation	Home, community or delivery team setting	Cross-sectional / longitudinal	Thematic analysis	FEQ 1: To what extent were the critical components of the intervention received by the target population? FEQ 2: How did inputs contribute to the intervention functioning? FEQ 5: To what extent was the intervention delivered as intended?
< 18 participants	Treatment group	Observation	Home, community or delivery team setting	Cross-sectional / longitudinal	Thematic analysis	FEQ 1: To what extent were the critical components of the intervention received by the target population? FEQ 2: How did inputs contribute to the intervention functioning?

² Adapted from Table 8 of the ACF1 summative evaluation protocol's 'IPE methods overview'

						FEQ 5: To what extent was the intervention delivered as intended?
Project lead / delivery team	NA	Interview	Delivery team or community setting	Cross-sectional	Thematic analysis	FEQ 2: How did inputs contribute to the intervention functioning? FEQ 4: How did local context affect intervention delivery? FEQ 5: To what extent was the intervention delivered as intended? FEQ 6: How did complexity affect intervention delivery? FEQ 8: Why did proximal outcomes change? FEQ 9: What was learned from how the intervention was delivered?
						FEQ 5: To what extent was the intervention delivered as intended? FEQ 6: How did complexity affect intervention delivery? FEQ 8: Why did proximal outcomes change? FEQ 9: What was learned from how the intervention was delivered?
Stakeholders (statutory / voluntary)	NA	Interview	Delivery team or community setting	Cross-sectional	Thematic analysis	FEQ 2: How did inputs contribute to the intervention functioning? FEQ 4: How did local context affect intervention delivery? FEQ 5: To what extent was the intervention delivered as intended? FEQ 8: Why did proximal outcomes change? FEQ 9: What was learned from how the intervention was delivered?
						FEQ 5: To what extent was the intervention delivered as intended? FEQ 8: Why did proximal outcomes change? FEQ 9: What was learned from how the intervention was delivered?
Project lead / delivery team	NA	Observation	Delivery team or community setting	Cross-sectional	Thematic analysis	FEQ 4: How did local context affect intervention delivery? FEQ 5: To what extent was the intervention
						FEQ 5: To what extent was the intervention

						delivered as intended? FEQ 6: How did complexity affect intervention delivery?
Stakeholders (statutory / voluntary)	All treatment population will be sampled	Observation	Delivery team or community setting	Cross-sectional	Thematic analysis	FEQ 2: How did inputs contribute to the intervention functioning?
Routine admin data	All routine process data will be included	Routine data collection	N/A	Longitudinal	Descriptive statistics	FEQ 1: To what extent were the critical components of the intervention received by the target population? FEQ 2: How did inputs contribute to the intervention functioning? FEQ 3: Who did the intervention work for and how? FEQ 5: To what extent was the intervention delivered as intended? FEQ 6: How did complexity affect intervention delivery?

Appendix A

FEQ 1: To what extent were the three components of the intervention, as required by the YEF framework, received by the treatment population groups (Compliance)?

Aim: Early identification of whether project sites are meeting **identification**, **engagement** and **retention** goals. Participants from target populations are recruited and enrolled in the intervention and monitored using a participant tracking tool. Participant data may reveal challenges with, for example, identification, recruitment and referral processes and ability to retain participants.

Indicators of intervention reaching participants	sub-evaluation questions
Resourcing	What local resources were available to ensure the intervention reached the intended number of participants?
	How were local communities represented in the resourcing of the intervention?
	What were the local resourcing challenges affecting delivery?
	How was the intervention funded and resourced?
	Did you rely on 'in-kind' support to resource the delivery team? If so, how much?
	What are your plans for ensuring a sustainable resourcing model? How sustainable is your resourcing model?
Identification	How was the target population defined?
	What data were used to identify target population groups (i.e., those eligible for randomisation)

Indicators of intervention reaching participants	sub-evaluation questions
	What extent was the target population group representative of the wider population?
	How was race equity assessed and maintained in the identification of target population groups.
Engagement	How were participants recruited?
	What intervention component was used in their recruitment?
	What was the most successful approach in recruitment?
	Did the recruitment of participants follow the specified process in the local delivery plan?
	What was different (compared to BAU processes) about how young people affiliated to groups/violence were identified?
	What proportion of the eligible population had a link to group violence?
	What was the recruitment method for each individual?
	What was the distribution and combination of selection criteria for each individual?
	How was race equity assessed and maintained in the participant recruitment process?
	To what extent was the distribution of recruitment outcomes racially proportionate?

Indicators of intervention reaching participants	sub-evaluation questions
Retention	What was the distribution of follow-up engagement beyond individual contact?
	What was the pathway followed through the intervention (flowchart)
	How was race equity assessed and maintained in the retention of target individuals?
	To what extent were retention outcomes racially proportionate?
Inclusion	What was the individual-level data on inclusion, eligibility and nature of group violence?
	How was gang or group membership defined?
	How was the racial equity of gang or group membership assessed and maintained?
Contact	To what extent did contact with participants follow the specified process in their local delivery plan?
	What communication methods were used to initiate contact with participants (e.g., phone call, email, letter, in-person visit)?
	Who made initial contact with participants, e.g., police, navigators, etc.?
	What training did those initiating contact with participants receive for this role?
	How representative were these individuals in making initial contact with the participant group?

Indicators of intervention reaching participants	sub-evaluation questions
	How was the racial equity of the communication method assessed and maintained?
Referral	What proportion of initial contacts led to referrals into the programme for control or treatment?
	Where relevant to site programme, what referral pathways were available into the programme
	Where relevant to the site programme, what was the distribution of referral sources?
	What factors were used to assess risk (e.g., school exclusion, already-allocated social workers, etc.)?
	How were navigators or support workers used?
	How were community members used?
	How was consent to participate in the programme obtained?
	How was the racial equity of referral mechanisms and processes assessed and maintained?
	To what extent were referral outcomes racially proportionate?
Follow-up	What were the follow-up processes used to retain participants on the programme?

Indicators of intervention reaching participants	sub-evaluation questions
	What was the distribution of follow-up attempts with each individual?
	What was the criteria for deciding to follow up with participants?
	How was the racial equity of follow-up efforts assessed and maintained?
	To what extent were follow-up outcomes racially proportionate?

FEQ 2. How are inputs contributing to the intervention functioning?

FEQ 3: Who did the intervention work for and how?

Aim: Early identification of whether sufficient resources and input processes are available to implement the intervention at local project sites and of the beneficiaries of the intervention.

ToC input categories	Indicators	sub-evaluation questions	
Personnel	Sufficient staff to deliver intervention, e.g. programme manager, navigators, analysts, support services, enforcement, community, voluntary, statutory	How did you ensure that you had sufficient staffing resources? What did you do in the event of understaffing?	
	Staff recruitment	How were staff identified and recruited?	
			What was the value structure you set for the delivery team (e.g., neutrality, equity)?
			What skills were identified as essential to the delivery of the intervention?
			What skill gaps did you have?
			What structural process did you have to unblock or overcome to recruit staff?
	Staff training		How did you ensure you selected the right staff for the project?
			What were your training needs?
			What training was provided to the staff on the delivery team? What training was provided

ToC input categories	Indicators	sub-evaluation questions
		regarding race equity?
		What knowledge gaps did you have that impacted the delivery of the intervention?
		What gaps in knowledge were there in relation to race equity? How did these affect planning and delivery?
	Case management	How quickly did you identify suitable cases?
		What systems or processes were used to capture data about case management and exit from the programme?
		How were cases triaged or risk assessed? What methods or frameworks were used?
		Who was present during the triage/risk assessment discussions?
		What was the typical length of time between the initial offer of support and individual access to that support?
		What was the distribution of the caseload?
		Did you face any challenges with the volume of cases and case management? If so, what were they?
		How much time was allocated to each case per week, including face-to-face and admin?

ToC input categories	Indicators	sub-evaluation questions
		How were cases closed? What was the process?
		What was the average time each case was in the programme?
		How did participants exit the programme?
		What support was provided to participants upon exiting the programme?
	Supervisory structures	What were the supervisory structures to manage the progression of cases?
		How often were case supervisory meetings held? Who were these chaired by? Who was present?
		How did these structures change during the project?
	Size of core delivery teams	What was the structure of the core delivery team?
		How did the size of the delivery team change during the project?
	Information	Police intelligence
Partner intelligence		What processes did you have to gather partnership intelligence? Did you consider those processes to be effective?
Community intelligence		What processes did you have to gather community intelligence? Did you consider those processes to

ToC input categories	Indicators	sub-evaluation questions
		be effective?
	Analysis	How many analysts were in the delivery team?
		How were the analysts used?
		How did their role change during the project?
		How were all the sources of intelligence analysed and evaluated? By analyst or team? How was race equity built into these processes?
Referrals	Schools	How many referrals were from schools and education stakeholders?
	Police	How many referrals were from the police?
	Individual	How many voluntary referrals?
		How did these individuals access the self-referral process?
	Family	How many referrals were from family members or guardians?
	Trigger offence	How many referrals were from trigger offences?
	Statutory services	How many referrals were from statutory services?

ToC input categories	Indicators	sub-evaluation questions
	Voluntary services	How many referrals were from voluntary or third-sector services?
Services	Partnerships	What was the extent of your partnership agreements?
		How effective were local partnership agreements?
		How could these local agreements be improved?
	Availability when needed	Were partner agencies available when needed to manage or progress cases?
	Accessibility	How accessible were support services provided by partner agencies?
		How were support services aligned to the individual needs of participants? In what way was race equity considered in the provision or offer of these services?
What were the take-up rates? What were the completion rates, if applicable? What were the barriers?		
Funding	Sufficient for delivery	How did funding contribute to the delivery of the intervention?
		How much has been spent so far?
		How did you prioritise your funding allocation?

ToC input categories	Indicators	sub-evaluation questions
		How much was spent vs. forecast?
		How much of your core delivery team is funded 'in-kind'?
		In what areas could efficiency savings be made?
		How could the intervention be delivered at a reduced cost?
	Local funding support	What level of local funding support did you have?
	Central funding support	What level of central funding support did you have?
		How well did you manage costs?
What were the challenges and risks related to funding structures?		
Organisational	Leadership support	Did the project have robust organisational structures in place for support, escalation and challenge?
		What leadership support did the project have? Was this sufficient?
	Monitoring/oversight	What were the high-level governance and oversight arrangements? Were these effective?
		How did these structures change through the project?

ToC input categories	Indicators	sub-evaluation questions
	Data systems	What data systems did the project need?
	Data sharing/access	What data sharing requirements were required?
		What challenges were there in relation to data systems, sharing and access?
Space	Appropriate for intervention activities	How was 'space' given to deliver the intervention?
Time	Sufficient staff time to deliver intervention	Were staff given the appropriate time to deliver the intervention?
	Sufficient supervisory time to manage intervention	Were supervisors given the appropriate time to deliver the intervention?
Experience of intervention	How did the treated population describe the intervention?	To what extent did each of the following theories contribute to any change?
		1. Deterrence (General)
		2. Procedural Justice, Police Legitimacy, Trust
		3. Social Control (establishment of trusted relationships, positive socialisation, removal of antisocial peers)
		4. Reduction in Strain (e.g., overcoming social and economic deficits, gaining skills, experience, or qualifications)
		5. Differential Association (e.g., group disengagement, positive self-worth through

ToC input categories	Indicators	sub-evaluation questions
		socially acceptable metrics, replacement of group maintenance factors)
		6. Turning Points in Life

FEQ 4: How is local context and complexity affecting intervention delivery?

Aim: Early identification of whether conditions related to the local context have been identified sufficiently. Context is defined as relating to (i) individual capacities, (ii) interpersonal relationships, (iii) institutional settings, and (iv) wider structures and systems.

Level	Context	Factors	sub-evaluation questions
1	Individual capacities	Attitudes	What is the evidence of a shared value structure across the delivery team?
			What was the shared value structure?
			What are the attitudes of participants and local community groups towards the police?
			What are the attitudes of the community towards the necessity of the intervention?
			How was race equity maintained through the project?
		Knowledge	What knowledge gaps did the team have that affected delivery?
Skills	What skill gaps did the team have that affected delivery?		
2	Interpersonal relationships	Lines of communication	How effective was communication between stakeholders? How could this be improved?
		Administrative support	How effective was the administrative support? How could this be improved?
		Professional relationships	How effective were the professional relationships between stakeholders? How could these be improved?

Level	Context	Factors	sub-evaluation questions
		Contractual arrangements	What contractual arrangements were there between stakeholders for delivery? What terms were agreed upon? What were the associated challenges?
3	Institutional setting	Culture and norms	What evidence was there of a commitment to change behaviors or reduce youth violence in the delivery team?
			What evidence was there of a commitment to act beyond BAU in providing support and services to participants?
		Leadership	How open were senior leaders to the project?
			How open were senior leaders to innovation to reduce youth violence?
		Governance	How effective were governance and accountability structures? How could these be improved?
		4	Wider structures and systems
Where was the delivery team located (e.g., VRU, OPCC)?			
Availability of funding	How did the availability of funding affect the delivery of the intervention (e.g., YEF staged release)?		
	How effective were relationships with co-funders?		
Supporting policies	What local policies supported the delivery of the intervention?		
Competing policies/influences	What local policies opposed the delivery of the intervention?		

FEQ 5: Is the intervention being delivered as intended?

FEQ 6: How does complexity affect intervention delivery?

Aim: Early identification of whether the components of the intervention are being delivered according to the YEF FD framework and what changes are required to accommodate local context and population requirements.

- **[Fidelity/adherence]:** How has the intervention been delivered with fidelity/adherence to YEF guidance?
- **[Dosage]:** How much of the intended intervention has been delivered?
- **[Quality]:** How well were the different components of the intervention being delivered?
- **[Reach]:** What was the rate of participation by the intended population?
- **[Responsiveness]:** What extent did the participants engage with the intervention?
- **[Intervention differentiation]:** What extent were the intervention activities sufficiently different from existing practices?
- **[Adaptation]:** What changes were needed to accommodate context and population requirements?
- **[Complexity]:** How have the characteristics of complexity affected the delivery of the intervention?

Fidelity dimension	Description	sub-evaluation questions	Data source	Method	Measures
Adherence	The extent to which the intervention was delivered according to YEF guidance and Theory of Change	How was the intervention delivered, adhering to the YEF framework and local delivery plan?	Delivery team	Interviews Surveys Focus groups Observation Self-report tools	Methods for delivering intervention Activities for delivering intervention Content of intervention delivered to participants
		How much emphasis was placed on each element of the intervention (e.g., enforcement, community, support)?			
		How much did the delivered intervention adhere to the local Theory of Change?	Delivery team	Interviews Surveys Focus groups Observation Self-report tools	

Fidelity dimension	Description	sub-evaluation questions	Data source	Method	Measures
Dosage	The difference between the amount of intervention delivered and the amount that the intervention model recommended	How much of the intended intervention was delivered to participants?	Delivery team	Interviews Surveys Focus groups Observation Self-report tools	Number of interventions Frequency and duration Time spent delivering intervention
Quality	The expertise with which delivery teams implement intervention	How well were the different components of the intervention delivered?	Delivery team	Interviews Surveys Focus groups Observation Self-report tools	Implementer enthusiasm Preparedness Clarity of expression Interaction Responsiveness Use of relevant examples
Responsiveness	The way in which participants react to or engage with the intervention	To what extent did the participants engage with the intervention?	Delivery team	Interviews Surveys Focus groups Observation Self-report tools Routine data collection	Level of participation in intervention Level of interest in intervention by participants Perceptions about relevance and utility of intervention
		How satisfied were participants with the intervention?	Participants	Interviews Surveys Focus groups Self-report tools	
Differentiation	The degree to which the critical components of the intervention are distinguishable from each other and from other violence reduction programmes	To what extent were the intervention activities sufficiently different from each other?	Delivery team	Interviews Surveys Focus groups Observation Self-report tools	Degree in change from similar violence reduction interventions Rationale for difference in intervention across sites
		To what extent were the intervention activities and	Delivery team	Interviews Surveys	Perceptions about relevance of

Fidelity dimension	Description	sub-evaluation questions	Data source	Method	Measures
		Theory of Change sufficiently different across sites?		Focus groups	intervention in other violence reduction programmes
		To what extent were the intervention activities sufficiently different from other violence reduction programmes?	Delivery team	Interviews Surveys	
Adaptation	The type of changes needed to adapt the intervention to reflect local operating context and complexity	What changes were needed to accommodate the local context and population requirements?	Delivery team	Interviews Surveys Focus groups Self-report tools	Rationale for modifications to adapt to local context Content of intervention requiring change Perceptions as to why modifications were required
Complexity	The degree to which the characteristics and behaviours of complexity changed delivery of the intervention	What is the evidence for the characteristics and behaviours of complexity? How did the complexity affect the delivery of the intervention and the likelihood of proximal outcomes? How well were the challenges related to complexity understood during the implementation of the intervention? How were these challenges anticipated and managed during	Delivery team	Interviews Surveys Focus groups Observations Self-report tools	Non-linearity: Effects of inputs on outputs or outcomes are not proportional Adaptation: Learning in response to implementation of intervention Path dependency: Actions based on decisions that proceed them Emergence: Unexpected higher-level outcomes that cannot be predicted from lower-level actions Feedback: Positive feedback leads to acceleration in change to effect of intervention and outcomes

Fidelity dimension	Description	sub-evaluation questions	Data source	Method	Measures
		implementation of the intervention?			<p>Self-organisation: Higher-level outcomes occur due to autonomous decision-making at lower-levels</p> <p>Tipping point: Point where interventions Outcomes dramatically change</p>

Appendix B

Proximal outcome measures

FEQ 7.How are proximal outcomes changing?

FEQ 8.Why are proximal outcomes changing?

Aim: Early identification of whether proximal outcomes are changing and trending in the desired direction for the intervention, and what is known about why they are changing, i.e., what mechanisms are being activated and by whom.

Proximal (short-term) outcomes	Intended to increase or decrease	sub-evaluation questions (has it increased or decreased so far—is it better or worse?)	Benchmarks ³	Data sources
Increase in community confidence in local policing	Increase	To what degree has the intervention led to an increase in community confidence in local policing?	Baseline 1	CSEW; PCC routine and annual surveying data; Community impact assessments
Increase in community confidence in local statutory services	Increase	To what degree has the intervention led to an increase in community confidence in other local statutory services?	Baseline 2	PCC routine and annual surveying data; Community impact assessments
Increase in community and partnership working	Increase	To what degree has community and partnership working changed during the delivery of intervention?	Baseline 3	Prospective survey with delivery team
Increase in understanding of local violence patterns	Increase	To what degree has understanding of local violence patterns and dynamics increased during the delivery of the intervention?	Baseline 4	Police/LA intelligence inputs/outputs re. gang/group/individual violence
Increase in identification of young people at risk of	Increase	To what degree has understanding of local violence patterns and dynamics led to an	Baseline 5	Police/LA intelligence inputs and outputs gang/group/individual

³ Baseline indicators correspond to data points, if available, at start of early implementation phase

Proximal (short-term) outcomes	Intended to increase or decrease	sub-evaluation questions (has it increased or decreased so far—is it better or worse?)	Benchmarks ³	Data sources
violence		increase in identification of young people at risk of violence?		violence
Increase in referrals to intervention programme	Increase	To what degree has the number of referrals of participants within the population group increased?	Baseline 6	Routine police data
Enforcement methods applied to participants who refuse to modify behaviour	Increase	What percentage of the intervention population was targeted with enforcement methods?	Baseline 7	Routine process data
Decrease in local school exclusions	Decrease	To what degree has the intervention decreased local school exclusions in the population age group? PRU	Baseline 8	Annual exclusions data (18 months lag but might be able to get it from LAs)
Decrease in local hospital admissions	Decrease	To what extent has the intervention decreased local hospital admissions for violence in the population age group?	Baseline 9	Fingertips data (12m lag; A&E data may be available)
Decrease in violence participation	Decrease	To what extent has the intervention decreased participation in acts of violence in the population age group?	Baseline 10	Police.uk (1m lag); routine police data; geographical component
Decrease in violence victimisation	Decrease	To what extent has the intervention decreased violence victimisation in the population age group?	Baseline 11	Routine police data
Decrease in group violence	Decrease	To what extent has the intervention decreased group violence in the population group?	Baseline 12	Routine or bespoke police data
Increase in support services provided to participants	Increase	To what degree has the intervention increased the services offered to population groups?	Baseline 13	Routine data; LA or third sector data
Increase in data sharing	Increase	To what degree has the intervention increased?	Baseline 14	DSAs; routine data

Proximal (short-term) outcomes	Intended to increase or decrease	sub-evaluation questions (has it increased or decreased so far—is it better or worse?)	Benchmarks ³	Data sources
between stakeholders		data sharing between intervention stakeholders?		
Increase in number of arrest for violence	Increase	To what degree has the intervention increased the number of arrests for violent offenses?	Baseline 15	Routine police data



youthendowmentfund.org.uk



hello@youthendowmentfund.org.uk



[@YouthEndowFund](https://twitter.com/YouthEndowFund)

The Youth Endowment Fund Charitable Trust

Registered Charity Number: 1185413
