



THE  
LAB

# Ending Youth Violence Lab

---

An Evidence Manifesto

January 2024

Contributions  
from Dr Nick  
Axford and  
Dr Sajid  
Humayun.

Edited by  
Tom McBride  
and Jack  
Martin.





# Acknowledgements

This report has been written and edited by Tom McBride and Jack Martin, based on thinkpieces written by Nick Axford (University of Plymouth) and Sajid Humayun (University of Greenwich). We are very grateful to Nick and Sajid for their contributions, and for the support and guidance they continue to provide.

We are grateful to the Youth Endowment Fund and Stuart Roden for the funding which supports the work of the Lab.

## About

This report was first published in January 2024 and is available to download as a free PDF at: <https://www.bi.team/wp-content/uploads/2024/01/EYV-Lab-Evidence-Manifesto.pdf>

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



THE  
**BEHAVIOURAL  
INSIGHTS  
TEAM**



# Introduction

**Involvement in crime and violence has a devastating effect on the lives of children, their communities and wider society...**

In the year ending March 2022 there were 33,000 proven offences committed by children in England and Wales. Around 8,000 children were first time entrants to the criminal justice system, whilst just under 3,500 proven knife and offensive weapon offences were committed by children (Youth Justice Board for England and Wales, 2023).

Children caught up in serious violence have poor life prospects. On average they will gain fewer qualifications, earn less, suffer poorer mental and physical health, and die younger than their peers (Piquero et al., 2010; Piquero et al., 2014).

Supporting children to avoid serious violence is one of the most critical social policy problems we face. Unfortunately, despite good intentions, services and programmes won't always achieve the outcomes we strive for. Or, to put it simply, sometimes things just don't "work".

**...To tackle this problem, we need to know which services have the most impact.**

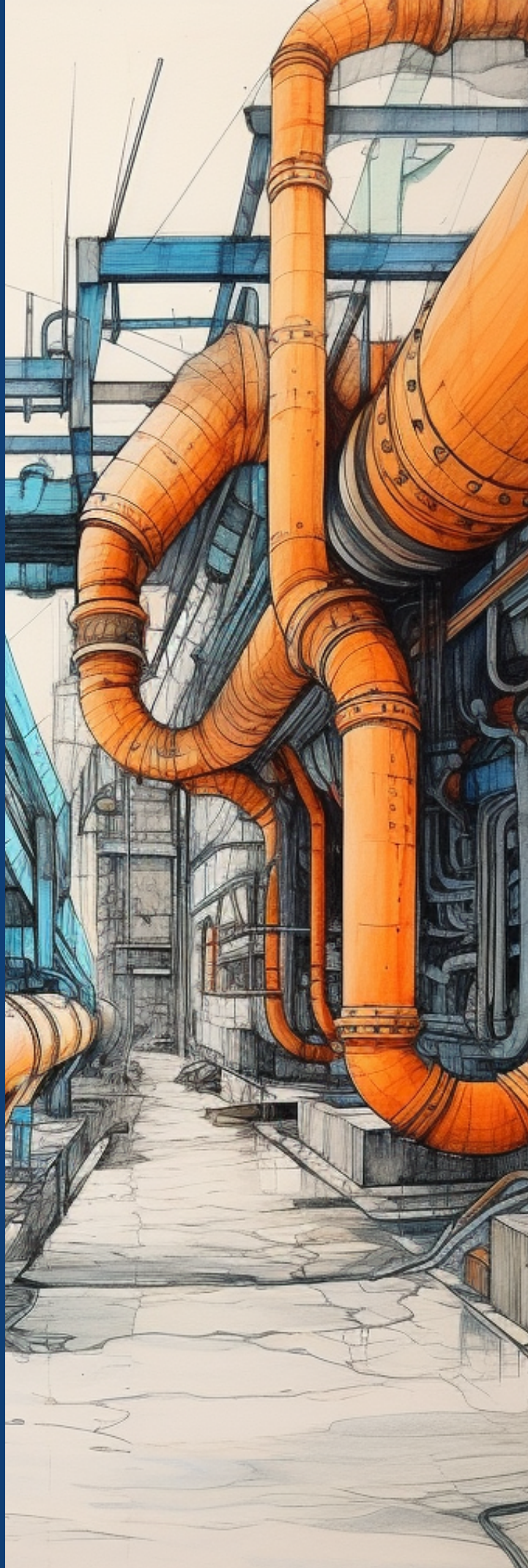
---

It's possible to identify what works and what doesn't, rather than simply hoping for the best. In the same way that medicine is rigorously tested in scientific trials before being used in practice, we can – and should – rigorously test and evaluate our public services.

Over the last 10 or so years there has been a real step-change in the UK in terms of the time and resource that has been directed at rigorously testing our services, and ensuring that proven and effective approaches are prioritised. We have come a long way, and we should celebrate the good work done by the [UK What Works Network](#) and others to enhance our understanding of how to address some of the most complex and pernicious challenges we face.

The [Youth Endowment Fund \(YEF\)](#) was established in 2019 to bring the What Works approach to reducing and preventing youth crime and violence. Since then, they have commissioned over 30 randomised control trials, with many more underway.

However, while many interventions and services show promise, far fewer are ready for this kind of evaluation. The [Ending Youth Violence Lab](#) was set up in 2022 with funding from YEF and philanthropist Stuart Roden. Our job is to work with and prepare services and interventions before YEF (or other organisations) rigorously test their impact, and to **build a strong pipeline of interventions.**





---

Eager to play our part in supporting and furthering the excellent work that has taken place over the last 10 years, we spoke to a number of experts. We wanted to learn from the experiences of the last decade, learn about how we could best deliver on our aims, and learn how to avoid the pitfalls involved in working with complex behavioural interventions.

Based on this we have set **five goals for the Ending Youth Violence Lab**. These build on our [strategy](#), and relate to ways we can **better design, refine and evaluate** services. If services are evaluated as well as possible, research can contribute even more to tackling the problem of youth violence.

We decided to publish these five goals for two reasons. The first is to share what we have learnt with the research and evaluation community - we think these goals should be shared by all of us who want to build the evidence on how to end youth violence.

The second is to publicly commit ourselves to best practice in social policy design and evaluation (hence 'manifesto') and to build on the efforts of others.



# Our 5 goals, at a glance

---

...for more detail on why and how we plan to action these goals, please see the following sections.

1

...to **minimise waste** in our evaluations.

2

...to **maximise learning** in our evaluations.

3

...to conduct a **thorough process of adaptation** before evaluating interventions developed overseas.

4

...to explore the **structural factors** which lead to youth violence, and test approaches which tackle them.

5

...to **build the evidence beyond discrete and branded interventions**.

# The first goal...

---

...to **minimise waste** in our evaluations.



## The context

There is pressure for organisations that deliver youth services to proceed to trial quickly, in order to demonstrate value for money and encourage future investment. This is driven partly by an accountability culture, and the belief that securing an endorsement from an external source - such as an evidence-based programme registry or toolkit - will increase the likelihood of the intervention being purchased by service commissioners.



## The issue

This context encourages moving too quickly and can lead to waste. It can incentivise progressing to impact evaluation *prematurely*, before important preparation has taken place. This means that service developers may spend more time and effort on trying to facilitate a trial ('proving'), rather than optimising their intervention and discerning the best and most feasible evaluation design for that intervention ('improving'). If they do go to trial prematurely, this risks finding no evidence of impact, which can be damaging for the provider (not to mention the opportunity cost). It also risks delivery issues such as poor fidelity, or high attrition, which may lead to inconclusive results and uninformative trials.

**There is a propensity for full-scale trials to happen too early in a programme's development.**

# The first goal...

---

...to **minimise waste** in our evaluations.

## The solution

It is important that the evaluation of early-stage interventions proceeds in careful steps, focusing on 'improving' initially before moving onto 'proving', and allowing the funder to change course if the approach does not suggest promise (Asmussen et al., 2019).

*Early-stage testing* is key, and typically involves two stages:



- The first is a **feasibility study**, which uses a mixed-methods approach to investigate whether an intervention can be delivered as intended, and recruit and retain its intended population. There is little point in proceeding to impact evaluation with an intervention which is insufficiently field-tested, too difficult or expensive to deliver, or where intended recipients do not engage with it.
- The second is a **pilot study**. These are small-scale randomised control trials. These studies are dress rehearsals for full-scale and rigorous trials, where the aim is not to demonstrate impact but to test out the feasibility of key evaluation procedures like randomisation and data collection. A pilot study gives us the opportunity to identify challenges quickly, early, and cheaply when the stakes are relatively low. Full-scale randomised control trials are burdensome and expensive - it's better to encounter issues in advance so they can be addressed in subsequent trials.

In addition, there should be **clear stop-go criteria** between these stages, where there is understanding of all parties that evaluation may not progress if the pilot identifies issues which can not be addressed in any future evaluation.

**The Lab will primarily focus on 'improving' rather than 'proving' through early-stage testing. This is a fundamental part of our strategy and how we will support funders of trials.**



## The second goal...

---

...to **maximise learning** in our evaluations.



### The context

Another consequence of the pressure to proceed to trial, and to demonstrate value quickly and cheaply, is a focus on **impact on outcomes at the expense of wider learning.**



### The issue

Establishing impact on outcomes is clearly important, and rightly the predominant aim of many forms of evaluation. However, trials often fail to maximise learning by only focusing on the difference between intervention and control arms on the outcomes of interest, with little or no additional data or analysis to help explain the results - and to indicate areas of potential improvement. These are what we might call 'thin trials', lean and efficient, but providing little information on how to proceed if a trial finds no impact, or on how to build on success if a trial has positive findings.

**There is a propensity for full-scale trials to miss key opportunities for learning.**

## The second goal...

---

...to **maximise learning** in our evaluations.

### The solution

'Thick' trials mitigate these issues by being more ambitious in terms of the amount and nature of data collected, and the type of analyses conducted (Axford et al., 2022). Thick trials can help facilitate further refinement of the intervention's design and may lead to ideas for improvements, by jettisoning elements that do not appear to be effective and allowing more focus on the key ingredients. They are often characterised by:

- *Mediator and moderator analyses* - which help with exploring what works for whom and why, and so understanding overall results and what might explain them (both positive, and negative). Small sample sizes may mean that this is only possible with descriptive analyses (O'Rourke & MacKinnon, 2018).
- *Qualitative research within trials* - which can help with explaining variation in outcomes, the mechanisms through which interventions have (or fail to have) impact, and why results might be disappointing, surprising or confusing (O'Cathain et al., 2014; Richards et al., 2019). This provides a richer picture of events, making trial results more informative.
- *Capturing what researchers have learnt from conducting the evaluation (and working with key stakeholders)* - Often information learnt through the many discussions and negotiations involved in evaluation (which is often where the significant barriers to a trial emerge) is not captured adequately. Capturing this information likely has to occur outside of formal data processes (like interviews or focus groups with practitioners), and instead through narratives provided by researchers.



**The Lab will focus on maximising learning by always conducting moderator analyses, and by embedding qualitative research in all evaluations.**

## The third goal...

---

...to conduct a **thorough process of adaptation** before evaluating interventions developed overseas.



### The context

Importing programmes that have been found to be effective in rigorous trials conducted *overseas* is common. However, when trialled in new settings, they are often found to have no effect. This is known as a failure to 'replicate'. Examples of replication failure of US interventions trialled in Europe include flagship violence and offending prevention interventions such as Functional Family Therapy (Humayun et al. 2017) and Multisystemic Therapy (Fonagy et al., 2018), as well as more upstream interventions such as the school-based social-emotional learning programme PATHS (Berry et al., 2016).

### The issue

We do not know exactly why this is, as there is no overarching analysis of the problem, but plausible theories include a failure to sufficiently adapt an intervention for its new context, factors relating to the context of the trial and differences in how the original and replication trials were conducted (Green et al., 2023).



Adaptation refers to a process of thoughtful and deliberate alteration of the design or delivery of an intervention to improve its fit or effectiveness in a new context. Whereas strict fidelity to intervention blueprints was once deemed essential to replication effectiveness, it is now recognised that reality is more complex and some adaptation of the intervention to the new context is likely to be needed (Green et al., 2023). The **failure to adapt well** is often cited as a potential reason for replication failure, in particular the risk that unplanned, excessive or insufficient changes could undermine the effectiveness of the programme and its fit with the local context.

**Many programmes transported from overseas to the UK are found to have little or no impact. We haven't learnt enough from these failures.**

## The third goal...

---

...to conduct a **thorough process of adaptation** before evaluating interventions developed overseas.

In terms of **trial factors and trial context**, it is quite likely that some of the factors that adversely affect trials attempting to replicate overseas are not fully reported or known (and go beyond insufficient cultural adaptation, and include trial flaws). Key ones to consider include:



- *The counterfactual* - One of the most likely explanations for replication failure are differences in the counterfactual. Trials will only answer the question 'Is the intervention being evaluated more effective than what we compared it to?'. So, the nature and effectiveness of the counterfactual itself is critical to trial outcomes. As a result, it is important to consider whether failures to replicate simply reflect superior business-as-usual services in the UK.
- *Factors affecting recruitment and differences in sample characteristics* - Another potential explanation for replication failure is that the sample recruited for the UK trial had significantly different characteristics to participants in previous studies, and the UK participants were less likely to benefit from the intervention. One explanation for why this occurs is evaluators and delivery bodies relaxing eligibility criteria in response to difficulties in recruiting an adequate number of participants.

**The explanation for this may include insufficient adaptation of the intervention, but may also include flaws with trials in the UK or contextual factors.**

## The third goal...

---

...to conduct a **thorough process of adaptation** before evaluating interventions developed overseas.

### The solution

There are at least three things we can do to mitigate these challenges:

- *Adapting interventions carefully* - There is much consensus about key principles for strong adaptation (e.g. involving diverse stakeholders, agreeing a way of working, protecting the theory of change and core components) and key steps (e.g. understanding the intervention and new community, consulting stakeholders, agreeing and making changes). Key considerations encouraged by most approaches include distinguishing between deep and surface adaptations, attending to both intervention and context (e.g. geographical, economic, organisational), and making pre-implementation but also ongoing changes during implementation.
- *Learning about the counterfactual* - It is important to conduct scoping work as part of trials to learn more about what business-as-usual services look like. This will give us a greater understanding of why replication may fail (or be less successful than hoped) and an opportunity to learn and share lessons.
- *Delivering to similar populations* - Unless there is a compelling reason to do otherwise, great care should be taken to ensure that eligibility criteria match previous successful studies, ensuring those who receive the intervention are those most likely to benefit. When implementing in a new setting, it is important to prioritise achieving the right sample over achieving a larger sample, and to understand the risks of relaxing eligibility.



**We will never deliver a programme from overseas without conducting careful adaptation work. We will collect detailed information on our samples, and business-as-usual services, to contextualise our findings and to learn and share lessons about implementing in a new context.**

## The fourth goal...

---

...to explore the **structural factors** which lead to youth violence, and test approaches which tackle them.



### The context

Low income and poverty are often the context but rarely the focus of prevention or early intervention programmes seeking to improve child and youth psychosocial outcomes (including offending, violence and anti-social behaviour). In other words, programmes often target low-income families or neighbourhoods, but are rarely overtly concerned with improving family economic well-being as an outcome (Axford & Berry, 2023).

### The issue

This situation is not sustainable. We know that low income and poverty have adverse causal effects on child and youth outcomes (e.g. physical and mental health, behaviour, educational performance) (Cooper & Stewart, 2021). We also know that improving family economic well-being can contribute to reducing adverse childhood experiences (e.g. domestic abuse, child maltreatment) and improving child and youth health and developmental outcomes (Cooper & Stewart, 2021; Courtin et al., 2019). There is also an ethical case for such intervention based on principles of equity, beneficence and not causing harm (for example, in the context of parenting interventions, by expecting parents/caregivers to change their parenting practices when their material living situation is challenging) (Axford & Berry, 2023).



Furthermore, lower socio-economic status is associated with lower rates of attendance in psychosocial interventions, whether owing to practical constraints (e.g. lack of transport, childcare) or feelings of stigma and shame, which in turn is likely to undermine their effectiveness (Berry et al., 2022).

**Low income/poverty are often the context but rarely the focus of prevention and early intervention programmes.**

## The fourth goal...

---

...to explore the **structural factors** which lead to youth violence, and test approaches which tackle them.

### The solution

There are at least two ways of mitigating this issue:

- *Funding and evaluating services that directly address low income and poverty* - Efforts to prevent or reduce youth violence and its precursors need to pay greater attention to family economic well-being in the context of systems and interventions designed to improve psychosocial outcomes. Clearly, policy change is a crucial part of this (e.g. minimum and living wages, tax and benefits system, employment programmes). However, there is also space for local initiatives to make an important difference here (e.g. support for female employment at a local authority level, poverty-proofing the school day) and innovation in frontline practice (e.g. better integrating income maximisation and debt counselling into psychosocial interventions, and promoting more poverty-aware or 'poverty-informed' practice among practitioners) (Axford & Berry, 2023).
- *Conducting evaluation that enables a better understanding of the influence of low income and poverty* - The impact that family economic status has on intervention attendance, engagement, and effectiveness should be examined more (i.e. explored as a moderating variable). It is also important to attempt to understand the mechanism explaining the connection between poverty and violence, given that many factors that may contribute towards violence overlap with poverty.



**We will always measure socioeconomic status and explore this as a moderator of programme impact, attendance and engagement. We are actively developing projects to fund and deliver activity that explicitly addresses these factors.**

## The fifth goal...

---

...to build the evidence beyond discrete and branded interventions.



### The context

Evaluation efforts in youth violence prevention and early intervention are predominantly focused on programmes or interventions - these are time-bound, manualised and specific packages of activity. They are rarely focused on *elements* of programmes - the activities and techniques used within programmes and interventions. Elements tend to fall into two categories – i) characteristics of interventions (e.g. duration intensity, delivery mode, targeting) and ii) activities, practices or ‘units of behavioural influence’ (e.g. skill-building, problem-solving, time out).

### The issue

An exclusive focus on programmes or interventions may:

- *Produce research which is less relevant and impactful* - Programmes/interventions are sometimes seen as a poor fit for local systems as they can be expensive, long or insufficiently flexible, or require specific workforces. Research suggests formal programmes/interventions make up a relatively small fraction of regular practice with children and families (Axford & Morpeth, 2013).
- *Inhibit our ability to learn* - If we evaluate interventions as a ‘black box’, we don’t learn which elements make them effective. This means we don’t know what to focus on when refining interventions or designing new interventions.



**Evaluation efforts tend to focus on programmes as a whole, rather than telling us which bits of them are most effective.**



## The fifth goal...

---

...to build the evidence beyond discrete and branded interventions.

### The solution

There are at least two ways of mitigating these issues, which include:

- *Mixed-method approaches that specifically attempt to investigate programme elements* - While it is difficult to test elements in isolation of the programme to which they belong, it is possible to conduct mixed-methods evaluation to come to a view on the relative value of different elements. This may include qualitative interviews with participants and practitioners, to gauge their views on how different elements and sessions of the intervention worked and their usefulness. It may also include quantitative surveys gauging satisfaction with different elements of the programme, or quantitative measures of fidelity indicating that some parts of the programme are harder to deliver than others. By triangulating the various pieces of information arising from a 'thick' trial, it is possible to come to a view on the acceptability, feasibility and perceptions of efficacy of different elements of a programme.
- *Comparing different versions of the same programme* - While it adds complexity and expense to deliver and evaluate multiple versions of the same intervention (i.e. one version with the element of interest, one version without), this can provide useful insight. This may be particularly viable in early-stage feasibility studies, where delivery is occurring on a smaller-scale.



**The Lab will attempt to learn as much as possible about the acceptability, feasibility and efficacy of programme elements through its evaluations.**



## Next steps

We have embedded these goals in the designs of our first set of evaluations (which you can read about [here](#)). Our [evaluation of GenPMTO](#), a parenting programme, exemplifies the goal of adapting carefully when transporting interventions from overseas. All three of our initial projects involve measuring SES/family income, conducting an early-stage study appropriate for the programme's level of development, and conducting 'thick' studies aiming to come to conclusions on the basis of triangulating several sources and types of data.

Looking forward, we will use these goals, alongside our strategy, to guide our work. For instance, we are pursuing multiple potential projects around investigating the impact of income boosting interventions on youth violence, and we are keen to explore projects that build the evidence for activity beyond discrete, branded interventions.

In the future, we look forward to working collaboratively with others in the sector to tackle youth violence, and to sharing our reflections on the experience of pursuing these goals in our work, and our successes, failures, and lessons learnt along the way.

# Bibliography

Asmussen et al. (2019). 10 steps for evaluation success. Early Intervention Foundation.

Axford, N., & Morpeth, L. (2013). Evidence-based programs in children's services: a critical appraisal. *Children and Youth Services Review*, 35(1), 268-277.

Axford, N., Berry, V., Lloyd, J., & Wyatt, K. (2022). How can we optimise learning from trials in child and adolescent mental health? *BMJ Evidence-Based Mental Health*, 25, 93-95.

Axford, N., & Berry, V. (2023). Money matters: time for prevention and early intervention to address family economic circumstances. *Journal of Prevention*, 44(3), 267-276.

Berry, V., Axford, N., Blower, S., Taylor, R. S., Edwards, R. T., Tobin, K., et al. (2016). The effectiveness and micro-costing analysis of a universal, school-based, social-emotional learning programme in the UK: a cluster-randomised controlled trial. *School Mental Health*, 8(2), 238-256.

Cooper, K., & Stewart, K. (2021). Does household income affect children's outcomes? A systematic review of the evidence. *Child Indicators Research*, 14, 981-1005.

Courtin, E., Allchin, E., Ding, A. J., & Layte, R. (2019). The role of socioeconomic interventions in reducing exposure to adverse childhood experiences: a systematic review. *Current Epidemiology Reports*, 6, 423-441.

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

Gardner, F., et al. (2023). Does social and economic disadvantage predict lower engagement with parenting support programs? An integrative analysis using individual participant data. *Prevention Science* 24(8), 1447-1458..





Green, F., Axford, N., Eastmond, N., Berry, V., Mannes, J., Allen, K., et al. (2023). Transporting an evidence-based youth development program to a new country: a narrative description and analysis of pre-implementation adaptation. *Journal of Prevention*, 44(6), 729-747.

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

O’Cathain, A., Good, J., Drabble, S., Thomas, K. J., Rudolph, A., & Hewison, J. (2014). Getting added value from using qualitative research with randomized controlled trials: a qualitative interview study. *Trials*, 15:215.

O’Rourke, H. P., & MacKinnon, D. P. (2018). Reasons for testing mediation in the absence of an intervention effect: a research imperative in prevention and intervention research. *J Stud Alcohol Drugs*, 79(2):171-181.

Piquero, A. R., Farrington, D. P., Nagin, D. S., & Moffitt, T. E. (2010). Trajectories of offending and their relation to life failure in late middle age: Findings from the Cambridge Study in Delinquent Development. *Journal of research in crime and delinquency*, 47(2), 151-173.

Piquero, A. R., Farrington, D. P., Shepherd, J. P., & Auty, K. (2014). Offending and early death in the Cambridge Study in Delinquent Development. *Justice Quarterly*, 31(3), 445-472.

Richards, D. A., Bazeley, P., Borglin, G., Craig, P., Emsley, R., Frost, J., et al. (2019). Integrating quantitative and qualitative data and findings when undertaking randomised controlled trials. *BMJ Open*, 9:e032081.

Youth Justice Board for England and Wales (2023). *Youth justice statistics: 2021 to 2022*.