Problem-Oriented Policing

Toolkit technical report

Ciara Keenan, Jaimee S. Mallion, Jennifer Hanratty, Sophie Hedges, Lily Strange, Teresa Geraghty, Rebecca Watson, Lauren Murphy, Karen Allen, Parady Baptiste, Frances Lyons, Caroline Coady, Jenny Towers

May 2025

About National Children's Bureau

This report has been produced by the National Children's Bureau on behalf of the Youth Endowment Fund. The National Children's Bureau works collaboratively across the issues affecting children to influence policy and get services working together to deliver a better childhood. They were commissioned by the Youth Endowment Fund (YEF) as their Toolkit Partner 2023-2026.

About Youth Endowment Fund

The Youth Endowment Fund's mission is to prevent children and young people becoming involved in violence. They do this by finding out what works and building a movement to put this knowledge into practice. The fund was established in March 2019 by children's charity Impetus, with a £200m endowment and ten-year mandate from the Home Office. For more information, please visit www.youthendowmentfund.org.uk.

Acknowledgements

Our deepest appreciation goes to all those who shared their insights and expertise to enhance the quality of this research. Special acknowledgement is due to Dr. Laura Knight, Head of Toolkit, and Dr. Chad Hemady, Senior Research Manager at the Youth Endowment Fund, for their diligent work and unwavering support throughout this project. Our sincere appreciation also goes to Professor James Thomas and the entire team at the EPPI-Centre, whose dedication and countless hours have been instrumental in the development of this work. Additionally, we acknowledge the invaluable support from our partners at The Race Equality Foundation, especially Tracey Bignall and Jabeer Butt OBE, who have consistently challenged our thinking by encouraging us to view our findings through the lens of children and young people experiencing discrimination and disadvantage.

Abstract/Plain Language Summary

This report examines the effectiveness of Problem-Oriented Policing (POP) in reducing children and young people's involvement in crime and violence. POP is a strategic approach that identifies and addresses the underlying causes of crime in specific areas, using problem-solving methods rather than traditional enforcement alone.

Key findings:

- Problem-Oriented Policing is associated with a 24% reduction in violent crime outcomes, based on a meta-analysis of 38 outcomes across 11 studies.
- Problem-Oriented Policing is also associated with a 21% reduction in broader crime-related outcomes, based on a meta-analysis of 16 studies and 81 measured outcomes.
- POP is most effective when targeting high-risk individuals, such as gang-involved young people, rather than broader community-level interventions.
- While POP can lead to a reduction in violence, its effectiveness varies depending on location, population, and implementation strategies. In particular, *where* an intervention occurs and *whom* it targets matters more than its duration.
- The intervention is potentially more successful at improving social and psychological well-being than at reducing crime in the long term, suggesting that additional support strategies, such as education or mentoring, may enhance effectiveness.
- Implementation challenges include data limitations, for example conducting appropriate and time sensitive analyses to identify areas for intervention; inconsistent application across locations; and difficulties in inter-agency collaboration, for example the different priorities of agencies involved, including police, schools, and other community organisations.

Fourteen studies provided evidence related to implementation, including 5 studies from the UK.

- High-quality evidence on who benefits most from POP is lacking, particularly regarding groups such as neurodiverse young people, care-experienced young people, and those with special educational needs and disabilities (SEND).
- The overall confidence in the evidence is moderate (3 out of 5). Most studies were rated as moderate or high quality, and while results show consistent reductions in crime and violence, variations in study quality and context limit certainty. Extensive analyses explain much of this variation, supporting a moderate level of confidence in the findings.

Conclusion

Problem-Oriented Policing offers a promising approach to reducing children and young people's involvement in violence and crime, but successful implementation requires strong data analysis, inter-agency collaboration, and tailored interventions. More high-quality research and evaluation are needed to fully understand its long-term impact and how it can be best adapted to different community needs.

Table of Contents

List of tables

List of figures

Objective and Approach

The objective of this report is to review the evidence on the effectiveness of Problem-Oriented Policing (POP) as a strategy to prevent violence involving children and young people. POP is a policing approach that focuses on identifying and addressing the root causes of crime through systematic problem-solving rather than traditional law enforcement methods alone. The goal is to develop targeted, evidence-based interventions that reduce crime and improve community safety, particularly for children and young people at risk of involvement in violence and crime.

This report is based on a comprehensive systematic review and meta-analysis of existing research on POP interventions. The evidence synthesis includes:

- 16 studies assessing the impact of POP on violence involving children and young people, and related outcomes.
- 81 measured outcomes, covering crime reduction, community engagement, and social well-being. 38 of these outcomes directly measured violent crime.
- Multiple study designs, including Randomised Controlled Trials (RCTs), Quasi-Experimental Designs (QEDs), and other quantitative evaluations.
- Implementation analysis of 14 studies, exploring how factors such as intervention duration, intensity, and target populations influence effectiveness.

By integrating quantitative meta-analysis with qualitative insights from implementation studies, this report provides a comprehensive evaluation of POP's impact, effectiveness, and practical considerations for policymakers, practitioners, and researchers.

The rest of this report is structured as follows: First, the **Description of the Intervention** outlines the key components of POP and its intended implementation. Second, **How Effective is the Intervention?** presents findings from our meta-analysis on crime reduction and broader social outcomes. Third, **Who Does it Work For?** examines evidence on the populations that benefit most

from POP. Fourth, **What Factors Affect Implementation?** explores key facilitators and barriers using Proctor's Implementation Outcome Framework. Fifth, **How Much Does It Cost?** reviews available cost data. Finally, the **Conclusion and Takeaway Messages** summarises key findings and recommendations, followed by **Appendices** detailing the systematic review methodology and characteristics of included research.

Description of the Intervention

Problem-Oriented Policing (POP; also known as Problem-Solving Policing) is a strategic approach that focuses on identifying and addressing the underlying causes of crime and disorder within specific areas. This method involves a systematic process known as the SARA model—Scanning, Analysis, Response, and Assessment—to develop tailored solutions for distinct issues. By emphasising problem-solving and preventive measures, POP aims to create sustainable reductions in crime by tackling its root causes (A. Braga, 2008a).

While both POP and hot spots policing[1] concentrate on areas with high crime rates, they are conceptually distinct (Modise, 2023). Hot spots policing primarily involves increasing police presence in designated high-crime areas to deter criminal activity through visible enforcement. In contrast, POP delves deeper into the specific problems contributing to crime in these areas, seeking to implement customised interventions that address these issues at their core. For instance, rather than solely increasing patrols in places with high levels of violence involving young people, POP would analyse factors such as environmental design, community engagement, and local socioeconomic conditions to develop a comprehensive response (Telep & Hibdon, 2019).

In the following section we provide detail on the interventions which inform this report, noting their key components, any equipment, materials, supplies or training required, the duration and intensity of interventions, who delivered the interventions, and where and how the interventions were delivered.

Intervention details

Features of the approach

Problem-oriented policing involves tailoring interventions to local problems, so each intervention is unique. However, many problem-oriented policing interventions to reduce violence involving young people share key components,

[1] https://youthendowmentfund.org.uk/toolkit/hot-spots-policing/

including the use of the SARA model, the identification and policing of 'hot spots', preventative work with communities and schools, and improvements to the local environment (Bullock et al., 2023).

The SARA model has four stages: Scanning, Analysis, Response, and Assessment (Eck & Spelman, 1987).[2] During the scanning phase, the problem to be addressed is identified and described. The analysis stage involves in-depth exploration of the underlying causes of the problem. Most interventions used crime data during the scanning and analysis stages (Baker & Wolfer, 2003; Boston Police Department, 2008; A. Braga et al., 2001; A. Braga, 2008b; A. Braga et al., 2008; A. Braga & Schnell, 2013; Gulf Breeze Police Department, 2008; Lancashire Constabulary, 2008; Transport for London, 2008), while some also sought community input through focus groups, surveys, and meetings (Boston Police Department, 2008; Maguire et al., 2003; Stokes et al., 1996). Some also used field observations and police officers' informal assessments (Boston Police Department, 2008; A. Braga & Schnell, 2013; Mazerolle et al., 1998). Scanning and analysis then inform the response stage, where tailored activities are developed and implemented to address the causes of the problem. Finally, the assessment stage measures the impact of the response to evaluate whether it had the intended effect. The evaluation methods used vary from simple pre/post comparisons exploring changes in crime rates over time (Boston Police Department, 2008; Braga, 2008b; Braga et al., 2001, 2008; Braga & Schnell, 2013; Gulf Breeze Police Department, 2008; Lancashire Constabulary, 2008; Maguire et al., 2003; Mazerolle et al., 1998; Transport for London, 2008), to more complex designs with control groups (Baker & Wolfer, 2003; Preston Early Intervention Partnership, 2012; Stokes et al., 1996), and even randomised controlled trials (Taylor et al., 2024; Weisburd et al., 2008).

POP interventions often focus on identifying specific areas at high risk for crime. Mapping software and crime statistics were frequently used to produce maps showing 'hot spots' for crime and violence, for example hot spots for anti-social behaviour and drug use (Baker & Wolfer, 2003), city streets with high levels of

[2] For more information on Problem-Oriented Policing and the SARA model, see the College of Policing's Crime reduction toolkit: https://www.college.police.uk/research/crime-reduction-toolkit/problem-oriented-policing

violent crime (Andrews, 2024; Beito & Sigler, 1997; A. Braga & Schnell, 2013; Taylor et al., 2024), streets around schools with high levels of violence against students (Stokes et al., 1996), bus stops or routes with high incidences of crime involving young people (Transport for London, 2008), and residential areas experiencing high crime rates (Beito & Sigler, 1997; Boston Police Department, 2008; Lancashire Constabulary, 2008; Mazerolle et al., 1998; Weisburd et al., 2008). This approach has similarities with hot spots policing, which also identifies locations where crime is most concentrated and focuses policing resources and activities on them.[3] Interventions included here differ due to their explicit use of the 'SARA' model and their focus on understanding and addressing the underlying problem.

POP can draw on a mixture of interventions and approaches to target the 'problem' of crime. As such, some POP interventions also incorporated features of hot spots policing and focused deterrence. For example, where interventions identified areas with high levels of crime, dedicated teams of police officers were deployed to carry out patrols in these areas with the aim of deterring crime (Andrews, 2024; Baker & Wolfer, 2003; Boston Police Department, 2008; Braga, 2008b; Braga et al., 2008; Braga et al., 2019;Braga & Schnell, 2013; Stokes et al., 1996; Taylor et al., 2024). Bicycle or foot patrol was a common feature of several interventions, bringing police officers closer to community members and enabling officers to build rapport and trust (Baker & Wolfer, 2003; A. Braga & Schnell, 2013; Gulf Breeze Police Department, 2008; Stokes et al., 1996). Similarly Interventions targeting gang violence drew on 'focused deterrence' or 'pulling levers' strategies as a problem-based response (Braga et al., 2001; Braga, 2008b; Braga et al., 2008).

Most of the POP interventions featured here involve elements of collaboration between the police and the local community or schools. Some interventions facilitated community meetings, Neighbourhood Watch groups, or presentations in schools to encourage community self-help and to provide information and awareness on crime prevention (Baker & Wolfer, 2003; Boston Police Department,

[3] For further information, see the YEF Toolkit entry for Hot spots policing at https://youthendowmentfund.org.uk/toolkit/hot-spots-policing/

2008; Beito & Sigler, 1997; Glover, 2002; Lancashire Constabulary, 2008; Weisburd et al., 2008; Stokes et al. 1996; Gulf Breeze Police Department, 2004; Braga et al., 2013; Transport for London, 2008). In Westwood, Massachusetts, police officers worked with staff and students at a local school to find solutions to bullying, threats and intimidation (Maguire et al., 2003). In Redlands, California, police officers also organised community recreational activities such as picnics, clean-ups, and parties to promote pro-social involvement and increase trust (Weisburd et al., 2008). In Preston, Lancashire, communities committed to neighbourhood clean-ups and young people who were at risk of offending were encouraged to join a community garden team (Lancashire Constabulary, 2008). Interventions targeting gang violence used community forums and meetings to communicate to gang members about the focused deterrence approach (Braga et al., 2001; Braga, 2008b; Braga et al., 2008).

As part of a more community-based strategy, many POP interventions use interagency working groups to link individuals with services to offer education and employment opportunities or to resolve health or housing issues, aiming to address some of the root causes of violence and offer alternatives to involvement in crime or gangs (Beito & Sigler, 1997; Braga et al., 2001; Braga et al., 2008; Braga, 2008b; Lancashire Constabulary, 2008; Mazerolle et al., 1998). The Nottingham Knife Crime Team and the Transport for London bus intervention both worked with Youth Offending Teams to provide support to young offenders (Andrews, 2024; Transport for London, 2008). As part of the 'Custody Experience' at Preston Police Station, young people at risk of offending were signposted to local activities and referred to social services or family support programmes where relevant (Preston Early Intervention Partnership, 2012).

Some interventions did use more traditional law enforcement approaches, often as part of a 'pulling levers' approach to target high-risk offenders. The Nottingham Knife Crime Team primarily used stop and search powers to deter and disrupt knife crime. The team often used road traffic or drug offences as a way to target known offenders and disrupt their activities (Andrews, 2024). In Jersey City, police officers enforced open warrants and used arrests and evictions

to remove high-risk offenders from the area (Mazerolle et al., 1998). In Stockton, Lowell, and Boston, interventions targeted gangs that committed violence, 'flooding' gang territories with police officers and probation officers to use all legal recourse to remove gang members from the streets or disrupt illegal activities (Braga et al., 2001; Braga, 2008b; Braga et al., 2008). In Preston, Lancashire, police targeted members of a local gang with Anti-Social Behaviour Orders (ASBOs) and secured convictions of high-profile offenders (Lancashire Constabulary, 2008). Also in Preston, the 'Custody Experience' intervention took inspiration from the US 'Scared Straight' programme, showing young people involved in anti-social behaviour around the custody complex, taking fingerprints and mugshots and explaining the consequences of arrest (Preston Early Intervention Partnership, 2012).

Finally, several interventions targeted the area itself, changing the environment to reduce opportunities for crime (known as 'target hardening'). Activities included removing overgrown vegetation, installing CCTV, repairing damaged fences, doors and windows, improving lighting, and using vandal-resistant materials (Baker & Wolfer, 2003; Boston Police Department, 2008; Braga & Schnell, 2013; Lancashire Constabulary, 2008; Mazerolle et al., 1998; Transport for London, 2008). In some cases, these measures may directly impact crime, for example removing vegetation that was screening criminal activities or repairing a broken window that could have been used to gain entry for burglary. According to 'broken windows theory', these measures may also impact on crime by sending the message that the community cares about the area and will not tolerate crime.

Equipment, materials or supplies required for implementation

The equipment required for POP interventions is dependent on the nature of the problem being addressed. However, analysis of the problem is always a key part of a POP approach, so data and analytical software, particularly mapping software, were used by most interventions (Andrews, 2024; Boston Police Department, 2008; Braga & Schnell, 2013; Stokes et al., 1996; Baker & Wolfer, 2003; Taylor et al., 2024; Weisburd et al., 2008). Andrews (2024) notes the importance of

data-connected laptops for the Nottingham Knife Crime Team in accessing up-to-date intelligence while patrolling.

Interventions involving police patrols required patrol cars or bicycles. The Nottingham Knife Crime Team had patrol cars with automatic number plate recognition technology, as well as devices to detect drugs during stops. The team were also equipped with Taser® conducted-energy devices and body-worn cameras, due to the high-risk nature of the intervention (Andrews, 2024).

Most interventions involved communication materials to ensure that the community was aware of the activities taking place and to raise awareness of the problem being addressed. Glover (2002) notes the importance of communication materials such as phones and copying services. In Boston, the police distributed door hangers and posters raising awareness of burglary risks and steps that residents could take to protect themselves (Boston Police Department, 2008). In Philadelphia, the school 'Safe Corridor' (areas patrolled by police to help CYP go to and from school safely) was advertised through an information pack given to students and parents (Stokes et al., 1996). In Redlands, community events were advertised through flyers and posters (Weisburd et al., 2008). Gang-focused 'pulling levers' interventions communicated the strong deterrence message through mass media communications including business cards, flyers, bus adverts, TV and radio announcements, and billboards (Braga et al., 2001; Braga, 2008b; Braga et al., 2008). Police in Preston implemented a media campaign via the local newspaper to promote positive neighbourhood action and reduce the fear of crime (Lancashire Constabulary, 2008). They also used the local newspaper to raise awareness of their 'Custody Experience' intervention, alongside a poster campaign with local organisations (Preston Early Intervention Partnership, 2012). Transport for London (2008) used posters to raise awareness of crime at bus stops and on public transport. Police in Gulf Breeze used a PowerPoint presentation on drug education in community education sessions (Gulf Breeze Police Department, 2004).

Training for intervention personnel

Around half of the interventions involved training for intervention personnel. Problem-oriented policing is a departure from traditional approaches to law enforcement and several interventions included training on POP and community policing (Braga & Schnell, 2013; Glover, 2002; Maguire et al., 2003; Taylor et al., 2024; Gulf Breeze Police Department, 2004). Some organisations also offered specific training on interactions with young people, for example bus drivers in London received training on addressing anti-social behaviour, and police officers taking part in a POP randomised controlled trial received training on positive police-youth interactions and youth development (Transport for London, 2008; Taylor et al., 2024).

Other projects offered bespoke training suited to the nature of the intervention. Police in the Nottingham Knife Crime Team received training in stop and search and behavioural analysis (Andrews, 2024). Police officers working on the Safe Corridor project were trained in using MapInfo software (Stokes et al., 1996).

Duration and intensity of the interventions

Most interventions did not state the duration or intensity of the intervention. Of the five which stated their duration, one lasted six weeks (Stokes et al., 1996), three lasted between one and two years (Maguire et al., 2003; Mazerolle et al., 1998; Taylor et al., 2024), and one lasted more than five years (Braga et al., 2008). None of the UK interventions reported their duration or intensity.

In terms of intensity, many interventions were ad hoc in response to need but some did state their frequency. The Nottingham Knife Crime Team was deployed on a fortnightly basis (Andrews, 2024). The school Safe Corridor was implemented daily during students' journeys to and from school (Stokes et al., 1996). Police also patrolled daily in the hot spots in the POP randomised controlled trial (Taylor et al., 2024).

Who delivered the interventions

All interventions were delivered by police officers. Many interventions also involved staff from other agencies, for example social workers, probation officers, prosecutors, parole officers, teachers, Youth Offending Teams, voluntary and community organisations, and faith leaders (Braga et al., 2001; Braga, 2008b; Braga et al., 2008; Braga & Schnell, 2013; Glover, 2002; Lancashire Constabulary, 2008; Preston Early Intervention Partnership, 2012; Maguire et al., 2003; Mazerolle et al., 1998; Weisburd et al., 2008; Transport for London, 2008; Andrews, 2024).

Where were the interventions delivered

The majority of the interventions took place in urban areas, generally in areas identified as being at high risk for crime or the territories of gangs committing violence (Braga et al., 2001; Braga, 2008b; Braga et al., 2008). Some interventions took place primarily on the street (Andrews, 2024; Braga & Schnell, 2013; Stokes et al., 1996; Taylor et al., 2024), others targeted other high-risk areas such as bus stops or parks associated with anti-social behaviour (Transport for London, 2008; Baker & Wolfer, 2003). Some interventions targeted residential housing areas (Beito & Sigler, 1997; Boston Police Department, 2008; Lancashire Constabulary, 2008; Mazerolle et al., 1998; Weisburd et al., 2008). Many interventions involved schools, for example police visiting schools to give presentations or patrolling the area around schools (Baker & Wolfer, 2003; Glover, 2002; Maguire et al., 2003; Stokes et al., 1996; Transport for London, 2008; Gulf Breeze Police Department, 2004). Gang forums were held at recreation centres, community centres, or schools (Braga et al., 2001; Braga, 2008b; Braga et al., 2008). The 'Custody Experience' intervention in Preston took place in the custody complex at Preston Police Station (Preston Early Intervention Partnership, 2012).

How were the interventions delivered

Interventions were delivered in different ways but all involved at least some face-to-face delivery by police officers through patrols and community meetings or activities. Many also included the delivery of messaging through targeted media channels, for example through poster or flyer campaigns, billboards, and TV and

radio announcements (Braga et al., 2001; Braga, 2008b; Braga et al., 2008; Braga & Schnell, 2013; Lancashire Constabulary, 2008; Preston Early Intervention Partnership, 2012; Stokes et al., 1996; Transport for London, 2008; Weisburd et al., 2008).

How Effective is the Intervention?

This section examines the effectiveness of POP in reducing violence and crime-related outcomes through a systematic review and meta-analysis, to provide a robust and objective summary of existing evidence, incorporating advanced statistical techniques, including robust variance estimators(Pustejovsky & Tipton, 2022), for improved accuracy. We present summary results from two separate meta-analyses below:

Table 1 Summary of findings on violence and crime outcomes

Outcome	Log RIRR	CI (95%)	*P*	% reduction	Impact rating	Number of studies	Evidence rating
Violence	-0.27	-0.36 to -0.19	< 0.0001 ***	-24%	Moderate	11	3
Crime & Offending	-0.26	-0.36 to -0.16	< 0.0001 ***	-23%	Moderate	14	3

Table 2 Summary of findings on other related outcomes

Outcome	Log RIRR	CI	P	% reduction	Number of studies
Community & Social Relationships	-0.16	-0.26 to -0.06	0.002 **	-15%[4]	5

[4] Negative log RIRR values indicate reductions in adverse outcomes, reflecting beneficial intervention effects. For outcomes like 'Community & Social Relationships', and 'Personal Development & Wellbeing' this denotes improved social connectedness, wellbeing and reduced victimisation. Outcome directions were standardised for consistency; see Appendix 1 (p. 77) for methodological details.

Drug and alcohol use	0.03	-0.02 to 0.08	0.28	+3%	1
Education	-0.02	-0.08 to 0.05	0.61	-2%	1
Personal Development & Wellbeing	-0.43	-0.93 to 0.06	0.08	-35%	1

Following these meta-analyses, meta-regressions were conducted to explore the effects of moderators including study design, study quality, intervention features, and outcomes.

See Appendix 1 for an overview of the methods used in this section and Appendix 3 for a list of the studies that provided data for the meta-analyses.

Meta-analysis of violence outcomes related to Problem-Oriented Policing

Problem-oriented policing is associated with a moderate effect, corresponding with a 24.0% reduction across violence outcomes, based on 38 measured outcomes across 11 studies. However, there is substantial variation in results.

The primary focus in the initial analysis is the reduction of violence, as defined by YEF. Violence is a broad construct that incorporates incidents/behaviours as well as convictable offences. Violence may be of a physical, verbal, psychological, or sexual nature (YEF, 2023: p.12). To assess this, relevant outcomes from the dataset were identified that align with this definition, including:

- Aggravated assault
- Rape and sexual assault
- Homicide
- Mugging (personal theft)

Eleven studies provided quantitative information across a variety of 38 violence-related outcomes related to the impact of a problem-oriented policing approach targeted at children and young people. These studies utilised a range of study designs including:

- RCT – Randomised Controlled Trials[5] (n=1)
- QED – Quasi-Experimental Designs[6] (n=9)
- PPD – Pre-Post designs[7] (n=1)

These 11 studies spanned multiple decades, with the earliest conducted by Stokes et al. (1996) and the most recent by Taylor et al. (2024). Most studies that measured violence were conducted in the US (n=10), and one in the UK.

Across the 38 measured outcomes, two outcome categories were identified within the YEF Outcomes Framework. These categories capture different aspects of violence involving young people and include:

1. Breaking the law or offending behaviour (k=32, 84.2%)
2. Being a victim of a violent crime (k=6, 15.8%)

The majority of violent outcomes were derived from crime statistics (k=32, 84.2%), with a smaller number being derived from self-completion questionnaires (k=5, 13.2%) and one-to-one interviews (k=1, 2.6%).

Authors employed various methodologies to assess the impact of problem - oriented policing on violence, including:

Multivariate and Hierarchical Models (k=22, 57.9%) such as Hierarchical linear modelling, Multivariate hierarchical count model, and Generalized linear models.

Pre-Post Comparative Analyses (k= 16, 42.1%) such as comparative analysis of reported incidents before and after intervention, percentage changes and reductions, and crime rate changes.

[5] RCTs randomly allocate participants to intervention and comparison groups
[6] QEDs use pre-existing or synthetic comparison groups rather than randomly allocating participants
[7] Single-group pre-post designs with only one post-intervention timepoint.

The studies in our analysis employed a mix of comparison types across different outcome measures, including:

- Comparison designs – Used for 24 outcomes (63.2%)
- Single-Series designs – Used for 14 outcomes (36.8%)

Results of the meta-analysis on violence outcomes

A total of $k = 38$ outcomes were included in the analysis. The observed outcomes ranged from -2.23 to 0.07 with the majority of estimates being negative (89%). The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.27$ $(95\%\ \text{CI:}\ -0.36\ \text{to}\ -0.19)$, therefore, the average outcome differed significantly from zero $z = -6.26, \quad p < 0.0001$. In practical terms, this means we have strong evidence[8] that the intervention has a real effect. **The Relative Incidence Rate Ratio (RIRR) is 0.77,** (95% CI: 0.70 to 0.83), with **a percentage reduction in violent outcomes of 24.0%**[9] for those receiving the intervention.

[8] Meaning it is very unlikely that the observed reduction in crime occurred by chance

[9] The IRR/RIRR can be converted to a percentage change metric using the following equation: % change=100 ×(IRR-1)

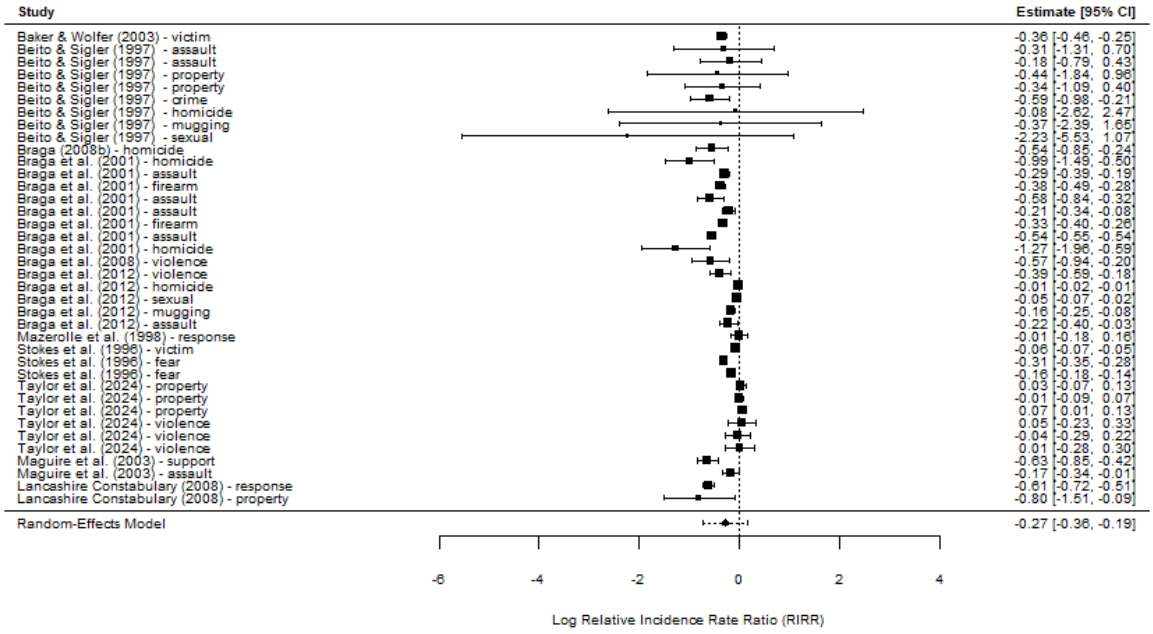


Figure 1 Forest plot showing the observed estimates of the random-effects model on violence outcomes.[10]

Next, we applied a Robust Variance Estimation (RVE; Hedges, et al. 2010) to adjust for within-study correlation when studies report multiple estimates[11]. This method accounts for potential dependencies in effect sizes due to:

- Repeated measures from the same study
- Multiple outcomes per study
- Clustered data (e.g., site-based interventions)

[10] As this is a place-based approach, some studies report composite outcomes related to sites where it is difficult to disentangle violence-specific incidents from other property-related offences. For example, (Taylor 2024) use a composite outcome measure that mix property-related and violence-related offences (i.e., (Site A) reports a composite of burglary, arson, shoplifting, and theft per month)). Sensitivity analyses were conducted to test the robustness of findings with and without these outcomes. The results remained stable, supporting the decision to retain them.

[11] More simply, when studies contribute multiple effect size estimates that are not completely independent

Table 3 RVE Output for meta-analysis on violence outcomes

	Estimate	SE	t-stat	d.f (Satt)	p-val (Satt)	Sig[12]
Intercept	-0.24[13]	0.099	-2.44	4.53	0.064	.

Results from RVE analysis:

Standard Error (SE): The initial analysis reported a standard error of 0.04. After applying RVE, the SE increased to 0.09 (see Table 3), suggesting that the original analysis may have underestimated within-study correlations, leading to a slight underestimation of the standard error. In other words, the original analysis likely underestimated the variability in the effect sizes, meaning the initial results may have given an overly precise estimate of the intervention's effect. The RVE correction provides a more realistic measure of uncertainty.

p-value: In the original analysis, the p-value was reported as highly significant, <0.0001. Following RVE adjustment, the p-value increased to 0.063 (Table 3**Error! Reference source not found.**). This change implies that the original analysis might have had inflated Type I error rates due to ignoring within-study dependencies[14]. This suggests that some of the statistical significance in the initial analysis may have been overstated due to the presence of correlated estimates.

In summary, the negative estimate (-0.24) provided by the RVE analysis still suggests a crime reduction effect associated with POP. **However, due to the increased standard error and higher p-value, we cannot be as confident in this result as we initially were**. Our application of RVE corrected for overconfidence in the original analysis, revealing that while POP may still be effective in reducing violence involving young people, the statistical certainty of this conclusion is

[12] p-values are annotated with significance codes where: *** denotes p-values less than 0.001; ** denotes p-values less than 0.01; * denotes p-values less than 0.05; . denotes p-values less than 0.1

[13] The intercept from the RVE model represents the average log Relative Incidence Rate Ratio (log RIRR) across the included studies

[14] This is important because ignoring these dependencies can lead to inflated statistical significance and an underestimation of variability.

weaker than initially suggested. This underscores the importance of accounting for study dependencies in meta-analyses to avoid overstating intervention effects.

Meta-analysis of all crime-related outcomes related to Problem-Oriented Policing

Problem-oriented policing is associated with a 23.0% reduction in Crime and Offending outcomes presented across 14 studies and a 21% reduction in broader crime-related outcomes across 81 measures from 16 studies; however, the effectiveness varies across different outcome categories.

Beyond violence reduction, we collected a wide range of additional outcomes to assess the impact of problem-oriented policing (POP) across various domains. These outcomes encompass multiple dimensions, including:

- Child-centred outcomes (e.g., behavioural change)
- Family and carer outcomes (e.g., parental stress)
- Peer and adult outcomes (e.g., beliefs about violence)
- School, professional, and community outcomes (e.g., collaboration improvements)
- Offending and crime outcomes (e.g., reoffending rates among children and young people)

Sixteen studies provided quantitative information across a variety of 81 effect sizes[15] related to the impact of a problem-oriented policing approach targeted at children and young people. These studies utilised a range of study designs including:

- Randomised Controlled Trials (n=2)

[15] In this review, an outcome refers to the specific variable measured to assess the effect of an intervention (e.g., crime rate, arrest frequency). An effect size is a statistical measure that quantifies the magnitude of the intervention's impact on an outcome. A single study may report multiple effect sizes if it assesses multiple outcomes, time points, or subgroups.

- Quasi-Experimental Designs (n=10)
- Other Quantitative Designs (n=1)
- Pre/Post Designs (n=3)

These 16 studies spanned multiple decades, with the earliest conducted by Stokes et al. (1996) and the most recent by Taylor et al. (2024). The majority of studies were conducted in US (n=13), with three conducted in the UK (Appendix 2).

Across the 81 measured outcomes, 11 outcome categories were identified within the YEF Outcomes Framework. These categories capture different aspects of behaviour, crime, and social factors and include:

3. Breaking the law or offending behaviour (n=13, k=45, 52.3%)
4. Community connectedness (n=3, k=10, 11.6%)
5. Victim of crime (n=3, k=10, 11.6%)
6. Family relationships and support (n=1, k=6, 7%)
7. Drug and alcohol use (n=1, k=4, 4.7%)
8. Positive and prosocial identity (n=1, k=3, 3.5%)
9. School engagement (n=1, k=2, 2.3%)
10. School environment (n=1, k=2, 2.3%)
11. Meaningful relationships (n=1, k=2, 2.3%)
12. Behavioural difficulties (externalising behaviours) (n=1, k=1, 1.2%)
13. Criminal peers (delinquent peers) (n=1, k=1, 1.2%)

The majority of outcomes were derived from Crime statistics (k=41, 47.7%); Self-completion questionnaires (k=34, 39.5%), One-to-one interviews (k=8, 9.3%) and Surveys (k=3, 3.5%). A list of these outcomes is available in Appendix 4. Authors employed various methodologies to assess the impact of problem -oriented policing, including:

Multivariate and Hierarchical Models (k=50, 58.1%) such as Hierarchical linear modelling, Multivariate hierarchical count model, and Generalized linear models.

Pre-Post Comparative Analyses (k=36, 41.9%) such as comparative analysis of reported incidents before and after intervention, percentage changes and reductions, and crime rate changes.

Results from the meta-analysis on crime and offending outcomes

A total of $k = 45$ outcomes were included in the analysis of crime and offending. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.26$ (95% CI: −0.36 to −0.16), indicating that the average outcome differed significantly from zero, $p < 0.0001$. In practical terms, this suggests strong evidence that the intervention has a real effect in reducing crime and offending behaviour. **The Relative Incidence Rate Ratio (RIRR) is 0.77** indicating a **23.0% reduction in crime and offending outcomes** for those receiving the intervention.

Results from the meta-analysis on all crime-related outcomes

A total of $k = 81$ outcomes were included in the analysis. The observed outcomes ranged from $-2.37 \text{ to } 0.10$, with the majority of estimates being negative (83%). The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.24 \quad (95\% \text{ CI: } -0.30 \text{ to } -0.17)$, therefore, the average outcome differed significantly from zero $z = -6.98, \quad p < 0.0001$. In practical terms, this means there is strong evidence[16] that the intervention has a real effect. **The Relative Incidence Rate Ratio (RIRR) is 0.80,** (95% CI: 0.74 to 0.84), with a **percentage reduction in crime-related outcomes of 21.0%** for those receiving the intervention.

A forest plot[17] showing the observed outcomes and the estimate based on the random-effects model is shown in **Error! Reference source not found.**.

[16] Meaning it is very unlikely that the observed reduction in crime occurred by chance

[17] Each study has been aggregated using RVE at the study-level for improved visualisation, thus, an overall mean effect is not displayed in the forest plot.

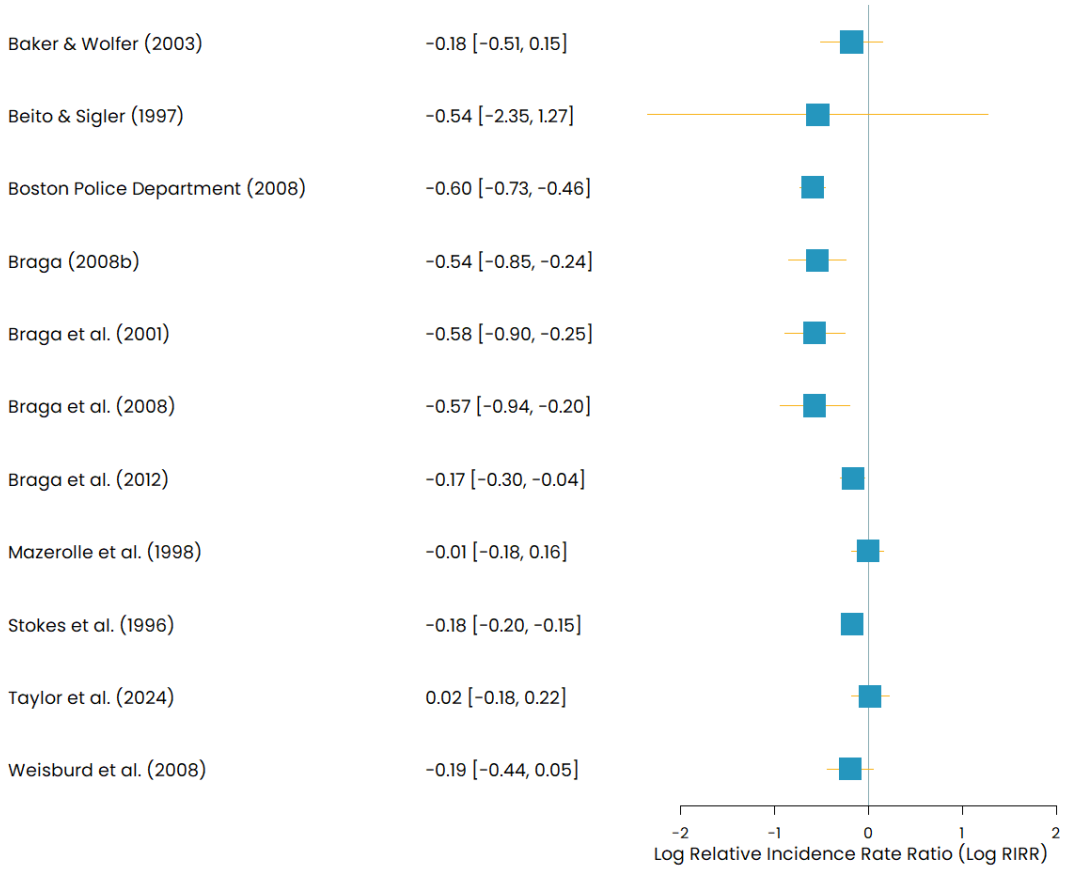


Figure 2 Forest plot showing the observed estimates of the random-effects model

According to the Q-test, the true outcomes appear to be heterogeneous $Q(80) = 20598.94, \quad p < 0.0001, \quad \widehat{\tau}^2 = 0.07, \quad I^2 = 99.4\%$. Such significant variability suggests that the differences in effect sizes are not solely due to random chance but may be influenced by specific study characteristics or contexts. To explore potential sources of this heterogeneity, conducting moderator analyses is recommended.

Publication Bias

To assess the potential of publication bias, we ran a few analyses. First, a funnel plot of the estimates is shown in Figure 3**Error! Reference source not found.**. Both the rank correlation and the regression test support this observed asymmetry $(p < 0.0001 \text{ and } p < 0.0001)$

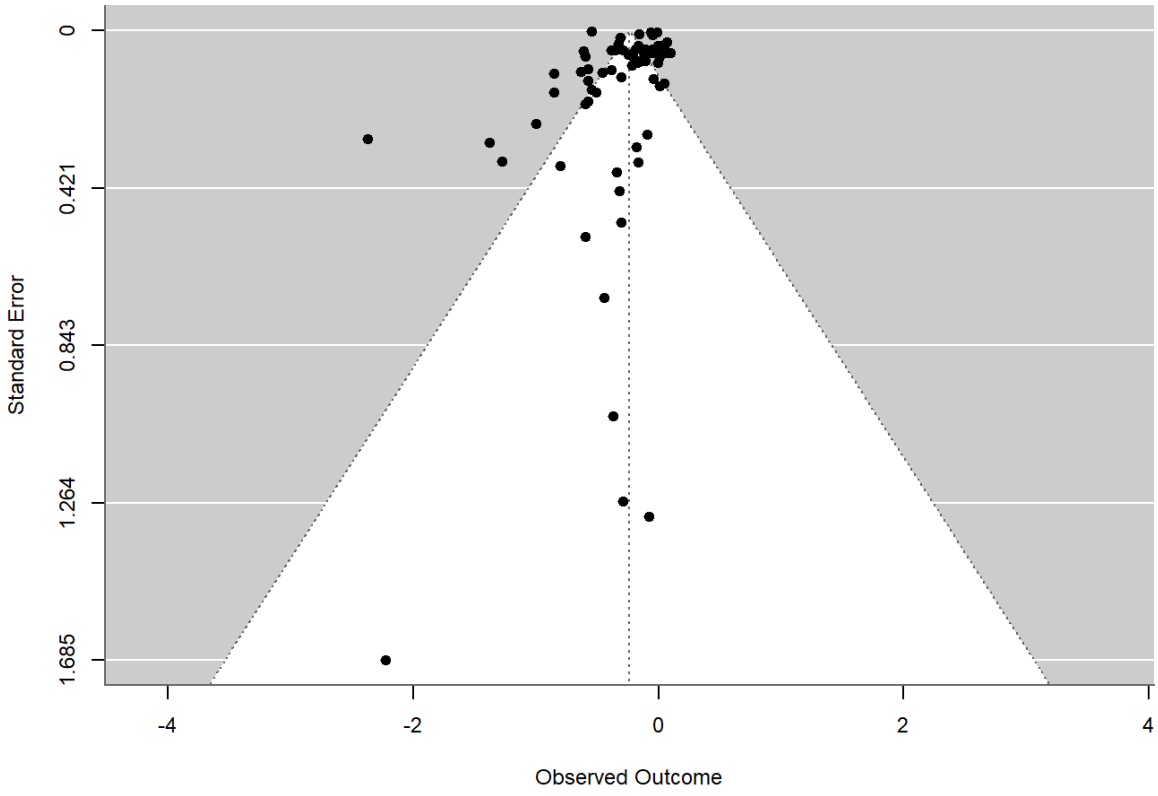


Figure 3 Funnel plot

The results of the funnel plot analysis suggest either the potential for publication bias and/or small-study effects, as indicated by the significant asymmetry detected through both the rank correlation test and regression test ($p < 0.0001$). This suggests that smaller studies reporting larger or more significant effects are more likely to be published, while smaller studies with null or negative results may be missing from the literature. As a result, the overall estimated intervention effect could be overstated. To address this, we used the trim and fill analysis which estimates and imputes missing studies to adjust for publication bias. The method estimated that 0 missing studies were needed to balance the funnel plot, suggesting that asymmetry due to publication bias is minimal. The adjusted pooled effect size remained unchanged at -0.24 (95% CI: -0.30 to -0.17, $p < 0.001$), consistent with the original meta-analysis results.

Sensitivity analysis

A 95% prediction interval for the true outcomes is given by -0.77 to 0.30 hence, although the average outcome is estimated to be negative (indicating a reduction in the targeted outcomes, such as violence involving children or young people, offending, or related risk factors), some studies may show positive estimates, suggesting an increase in these outcomes.

An examination of the studentized residuals revealed that one outcome had a value larger than ± 3.39 and may be a potential outlier in the context of this model. According to the Cook's distances, three further outcomes could be considered to be overly influential.

The team extracted these four outcomes from the file and checked the paper for accuracy, although these outcomes were larger than expected, they were included as they were as the author reported.

To maintain confidence in the robustness of the data, a leave-one-out sensitivity analysis was conducted to examine the influence of these effect sizes on the overall pooled effect size (-0.24). The results indicated that no single effect size significantly altered the overall effect estimate. The range of effect sizes remained stable when systematically omitting each of the 81 outcomes (ranging from -0.24 to -0.23), and statistical significance was maintained ($p < 0.001$ in all cases) suggesting that the results are robust. No study was identified as overly influential based on significant changes in the confidence interval or heterogeneity statistics. High heterogeneity was observed throughout the leave-one-out analysis (I^2 ~99.4%), indicating substantial between-study variability.

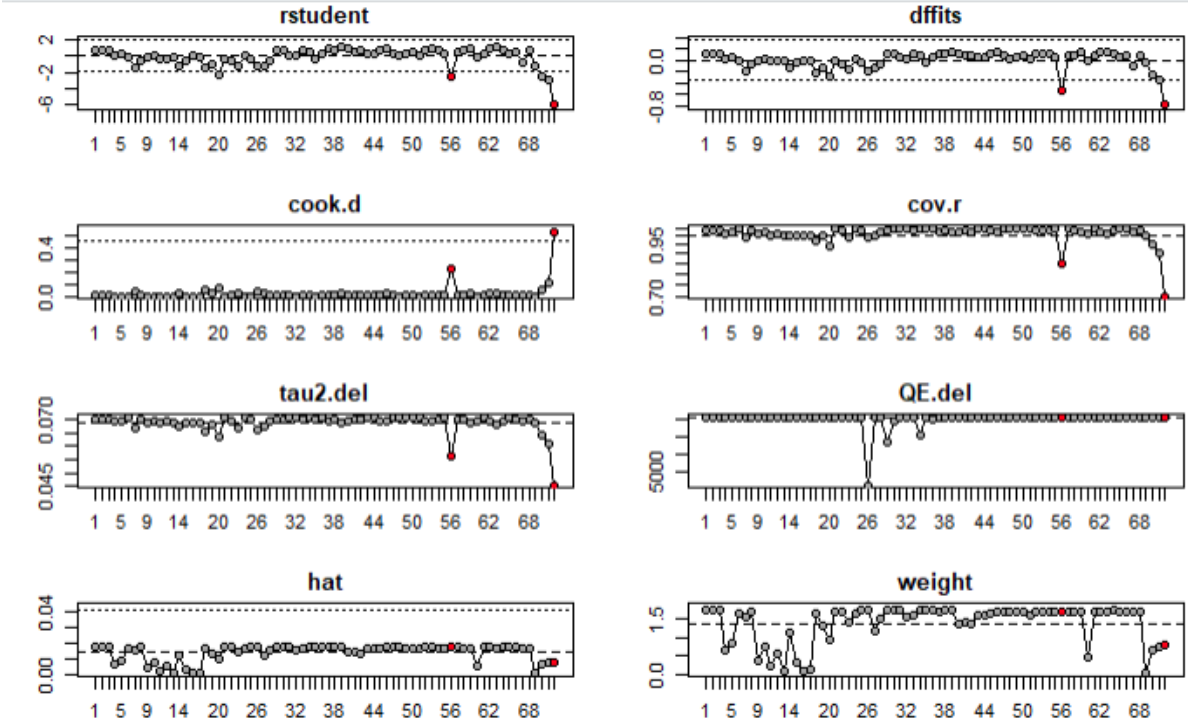


Figure 4 *Results from sensitivity analysis*

Robust variance estimation

RVE was then applied to adjust for within-study correlation when studies report multiple estimates.

Table 4 RVE Output for meta-analysis on all relevant outcomes

	Estimate	SE	t-stat	d.f (Satt)	p-val (Satt)	Sig
Intercept	-0.24	0.09	-2.44	4.53	0.06	.

Results from RVE analysis:

Standard Error (SE): The initial analysis reported a standard error of 0.03. After applying RVE, the SE increased to 0.09 (see Table 4**Error! Reference source not found.**), indicating that the original analysis likely underestimated variability in effect sizes.

p-value: In the original analysis, the p-value was reported as <0.001, indicating statistical significance. Following RVE adjustment, the p-value increased to 0.06 (**Error! Reference source not found.**). This change implies that the original analysis might have had inflated Type I error rates due to ignoring within-study dependencies.

The negative estimate (-0.24) still suggests a reduction in all relevant outcomes associated with POP but the level of certainty in this conclusion is lower than initially indicated.

Moderator analysis (Meta-regression models 1-5)

To explore the high sources of heterogeneity observed in our meta-analysis, we conducted five meta-regression analyses. This statistical approach allows us to examine how study-level characteristics, such as study design, intervention duration, and outcome measures, influence the effect sizes reported across studies. By incorporating these moderators into our model, we aim to identify factors contributing to variability in intervention effectiveness and gain a deeper understanding of the conditions under which problem-oriented policing is most effective.

In interpreting these results, it is crucial to note that in meta-analyses using log Relative Incidence Rate Ratios (log RIRR), moderators with negative coefficients indicate greater effectiveness in reducing crime outcomes. This is because a negative log RIRR corresponds to a reduction in crime rates, so moderators that further decrease the log RIRR enhance intervention effectiveness. Detailed methodological information is available in Appendix 1.

An important note is that with meta-analysis using Log RIRR, moderators with negative coefficients report **greater** effectiveness in reducing crime outcomes.

Moderator analysis 1. Study-Level Moderators

Characteristics added to model 1 include:

- Type of Publication
- Quality appraisal as assessed by the YEF-EQA tool
- Conflict of interest

Table 5 Results from moderator analysis 1 with study-level moderators

Moderator	Estimate	SE	Z-value	P-value	95% CI Lower	95% CI Upper	Sig
Intercept, β [18]	-0.60	0.23	-2.57	0.010	-1.05	-0.14	*
Publication Type: Journal Article	0.39	0.24	1.58	0.114	-0.09	0.86	
Publication Type: Research Report	0.42	0.26	1.58	0.115	-0.10	0.94	
YEF_EQA: Moderate	0.01	0.08	0.15	0.878	-0.15	0.18	
COI: Stated "Due to Competing Interests"	0.22	0.11	2.08	0.038	0.01	0.42	.

The meta-regression analysis for study-level moderators indicates that, while the baseline effect of the intervention is a significant reduction in crime-related outcomes (Intercept: β = -0.60, p = 0.01), studies declaring a conflict of interest reported **less** reduction in crime than those without such declarations (β = 0.22, p = 0.038, Table 5).

[18] The intercept represents the estimated effect size when all moderators that have been included in a model are at their reference levels. This serves as a baseline against which the effects of other moderator variables are compared.

The test of moderators (QM(df = 4) = 7.73, p-val = 0.01) indicates that the included moderators collectively explain a small portion of the variability in effect sizes, and model 1 accounts for 9.1% of observed heterogeneity.

Moderator analysis 2. Quality Moderators

Characteristics added to model 2 include:

- Study Design
- Quality appraisal as assessed by the YEF-EQA tool
- Comparison type

Table 6 Results from moderator analysis 2 with quality moderators

Moderator	Estimate	SE	Z-value	P-value	95% CI Lower	95% CI Upper	Sig
Intercept, β	-0.36	0.12	-3.05	0.002	-0.58	-0.13	**
Study Design: QED	0.20	0.12	1.70	0.089	-0.03	0.44	.
Study Design: RCT	0.18	0.08	2.11	0.035	0.01	0.35	*
YEF_EQA: Mod	0.09	0.14	0.64	0.523	-0.19	0.37	
YEF_EQA: Very Low	0.05	0.13	0.39	0.699	-0.21	0.31	
Comparison Type: SSRD	-0.16	0.10	-1.67	0.094	-0.35	0.03	.

The results indicate that study design significantly moderated effect sizes in this model. Specifically, studies employing Randomised Controlled Trials (RCTs) reported weaker reductions in crime compared to the reference group, with a

statistically significant positive coefficient (β = 0.18, p = 0.035). This suggests that more rigorous evaluation designs may yield more conservative estimates of effectiveness. Studies using Quasi-Experimental Designs (QEDs) also showed a trend toward smaller reductions in crime (β = 0.20, p = 0.089), although this was only marginally significant.

Study quality ratings (YEF-EQA) did not significantly moderate effect sizes. This suggests that reported effect sizes were not systematically influenced by the quality appraisal scores in this dataset.

The test of moderators (QM(df = 5) = 21.99, p = 0.0005) confirms that, collectively, study design, quality rating, and comparison type explain a statistically significant portion of the heterogeneity in effect sizes. The model accounts for 37.85% of the between-study variance, indicating that these study-level characteristics play a meaningful role in explaining variation in intervention effectiveness.

Moderator analysis 3. Intervention-Level Moderators.

Features of the Intervention added to model 3 include:

- Where the intervention happened
- The duration of the intervention
- The intensity of the intervention
- Whom the intervention was targeted at.

Table 7 *Results from moderator analysis 3 with intervention-level moderators*

Moderator	Estimate	SE	Z-value	P-value	95% CI Lower	95% CI Upper	Sig
Intercept, β	-0.64	0.38	-1.69	0.091	-1.39	0.10	.
Intervention Location: Educational Spaces	-0.07	0.32	-0.23	0.82	-0.70	0.56	

Intervention Location: Institutional	-0.02	0.31	-0.05	0.96	-0.62	0.59	
Intervention Location: Public and Recreational Spaces	0.42	0.19	2.16	0.031	0.04	0.80	*
Intervention Location: Urban Spaces	0.39	0.20	1.99	0.047	0.01	0.77	*
Duration: 1–3 Months	0.65	0.53	1.23	0.22	-0.39	1.69	
Duration: 2–3 Months	0.40	0.53	0.76	0.45	-0.63	1.44	
Duration: 3–6 Months	0.56	0.53	1.05	0.30	-0.48	1.59	
Duration: 1–2 Years	0.29	0.31	0.93	0.35	-0.32	0.90	
Duration: 2–3 Years	0.23	0.37	0.64	0.52	-0.48	0.95	
Duration: 3+ Years	0.30	0.37	0.83	0.41	-0.42	1.02	
Target Population: Community Members	-0.11	0.08	-1.33	0.18	-0.26	0.05	
Target Population: Gang-Involved CYP	-0.35	0.08	-4.63	<.0001	-0.50	-0.20	***

The meta-regression analysis assessing intervention characteristics found that location and target population significantly influenced intervention effectiveness. Interventions conducted in public and recreational spaces ($\beta = 0.42$, $p = 0.031$) and urban spaces ($\beta = 0.39$, $p = 0.047$) showed weaker reductions in crime compared to the baseline category (β), while interventions in educational spaces ($\beta = -0.07$, $p = 0.82$) did not show a statistically significant difference in crime reduction (Table 7).

Interventions targeting gang-involved children and young people ($\beta = -0.35$, $p < 0.0001$) showed the strongest crime reduction effects, suggesting that direct engagement with children and young people at higher risk of offending is particularly effective.

Intervention duration and intensity did not significantly predict effect size variation, meaning that how long an intervention lasted did not appear to impact its effectiveness in reducing crime.

Overall, this analysis highlights that **where an intervention occurs and whom it targets matters more than its duration.** Interventions targeting gang-involved young people show the strongest crime reduction effects, while interventions in public and recreational spaces and urban areas tend to report weaker reductions in crime. The model explains 61.3% of the observed heterogeneity ($R^2 = 61.3\%$), indicating that intervention characteristics play a substantial role in explaining differences in effectiveness, though some variability remains unexplained.

Moderator analysis 4. Outcome moderators

Features of the measured outcome added to model 4 include:

- YEF's Outcome Category
- The type of outcome (e.g., child-centred outcomes, school, professional, and community outcomes; and offending and crime outcomes)
- How was the data provided?
- What type of analysis was done to measure impact?

This meta-regression model evaluates how different outcome categories, outcome measurement methods, data sources, and types of analysis influence the effectiveness of POP. Model 4 accounts for 63.8% of the observed heterogeneity ($R^2 = 63.8\%$), highlighting the substantial role of these moderators in influencing effectiveness.

Table 8 *Results from moderator analysis 4 with outcome moderators*

Moderator	Estimate	SE	Z-value	P-value	95% CI Lower	95% CI Upper	Sig
Intercept, β	-1.38	0.34	-4.05	<.0001	-2.05	-0.71	***
Outcome Category: Breaking the Law or Offending	1.24	0.38	3.31	0.0009	0.51	1.98	***
Outcome Category: Community Connectedness	1.38	0.35	3.98	<.0001	0.70	2.06	***
Outcome Category: Criminal Peers	-0.99	0.48	-2.08	0.038	-1.92	-0.06	*
Outcome Category: Drug and Alcohol Use	1.41	0.35	4.01	<.0001	0.72	2.10	***
Outcome Category: Family Relationships and Support	1.30	0.35	3.73	0.0002	0.62	1.98	***
Outcome Category: Meaningful Relationships	1.33	0.36	3.69	0.0002	0.63	2.04	***
Outcome Category: Positive and Prosocial Identity	1.18	0.36	3.31	0.0009	0.48	1.87	***
Outcome Category: School Engagement	1.30	0.36	3.59	0.0003	0.59	2.01	***
Outcome Category: School Environment	1.39	0.36	3.84	0.0001	0.68	2.10	***
Outcome Category: Victim of Crime	1.33	0.35	3.80	0.0001	0.65	2.02	***

Outcome Measure: Offending and Crime Outcomes	-0.09	0.16	-0.58	0.56	-0.42	0.23	
Outcome Measure: School, Professional, and Community	-0.28	0.58	-0.48	0.63	-1.42	0.86	
Data Source: One-to-One Interview	0.44	0.58	0.76	0.45	-0.70	1.58	
Type of Analysis: Pre-Post Comparative Analyses	-0.32	0.07	-4.72	<.0001	-0.45	-0.18	***

The findings suggest that the type of outcome measured significantly influenced the reported effectiveness of POP interventions. Studies measuring outcomes related to criminal peers ($\beta = -0.99$, $p = 0.038$) reported stronger crime reduction effects than other categories, suggesting that interventions targeting peer influences were particularly effective.

Conversely, interventions targeting specific behaviours (e.g., drug use, family relationships, school engagement, and meaningful relationships) show weaker crime reduction effects compared to baseline (Table 8).

Additionally, the way data is collected or analysed does influence effect sizes, for example, pre-post comparative analyses ($\beta = -0.32$, $p < .0001$) showed significantly stronger reductions in crime, suggesting that studies using this approach may report larger effects compared to other methods.

Moderator analysis 5. Enhanced analysis of outcome moderators

The meta-regression analysis of outcome moderators (model 4 above) revealed much statistical significance – indicating some variables may be influencing POP's effectiveness, prompting a more in-depth investigation. To enhance the robustness and interpretability of our analysis, we categorised outcomes into higher-level groups, as detailed below:

Table 9 *Higher-level groupings of YEF outcome categories*

NCB new category	YEF Outcome Categories
Crime & Offending	• Breaking the law or offending behaviour
Community & Social Relationships	• Community connectedness • Family relationships and support • Meaningful relationships • Victim of crime
Education	• School engagement • School environment
Personal Development & Wellbeing	• Positive and prosocial identity • Behavioural difficulties (externalising behaviours)
Drug and alcohol use	• Drug and alcohol use

By consolidating outcomes into these broader categories, the aim is to increase the statistical power and relevance of the meta-regression analysis on additional outcomes (Moderator analysis 5), facilitating more meaningful interpretations of the intervention effects across diverse domains.

Table 10 *Results from moderator analysis 5 with new higher-level outcome groupings*

Moderator	Estimate	SE	Z-value	P-value	95% CI Lower	95% CI Upper	Sig
intercept	-0.17	0.052	-3.36	<0.001	-0.28	-0.073	***
Crime & Offending	-0.09	0.073	-1.22	0.22	-0.23	0.054	

Drug and alcohol use	0.20	0.14	1.51	0.13	-0.061	0.47	
Education	0.14	0.14	1.02	0.31	-0.13	0.40	
Personal Development & Wellbeing	-0.32	0.14	-2.18	0.029	-0.60	-0.032	*

The findings indicate that interventions targeting personal development and wellbeing (e.g., self-identity, behavioural improvements, and prosocial engagement) show significantly stronger crime reduction effects ($\beta = -0.32$, $p = 0.029$). This suggests that POP interventions may be particularly effective when they focus on improving social and psychological wellbeing, which in turn contributes to crime reduction.

Conversely, interventions targeting crime & offending behaviour ($\beta = -0.09$, $p = 0.22$), drug and alcohol use ($\beta = 0.20$, $p = 0.13$), and educational outcomes ($\beta = 0.14$, $p = 0.31$) did not show statistically significant effects. This suggests that while POP may have some impact in these areas, the effects are less pronounced or possibly require additional components to be effective (Table 10).

How Secure is the Evidence?

Violence outcomes

Our confidence in the findings on violence is moderate. The meta-analysis included 38 violence-related outcomes drawn from 11 studies assessing the impact of POP on children and young people.

Study quality, as assessed using the YEF-EQA, was moderate - high. Five studies were rated as high quality, and six were rated as moderate quality. Of the five high-quality studies, one was an RCT **(Type A)**, and the other four were QEDs **(Type B).** All of the moderate quality studies fell within **Type C impact evaluations** (four QEDs, and one PPD), resulting in an initial evidence security rating of **Level 3**.

While substantial heterogeneity is evident in the meta-analysis, reflected in the RVE adjustment and broad range of observed effects, subsequent moderator analyses account for much of this variation. As a result, the initial evidence security rating was not downgraded, and **an evidence security rating of 3 out of 5 is maintained.**

Crime and offending outcomes

Our confidence in the findings on crime and offending is moderate. The meta-analysis included 45 crime and offending outcomes drawn from 14 studies assessing the impact of POP on children and young people.

Study quality, as assessed using the YEF-EQA, was moderate - high. Four studies were rated as high quality, eight were rated as moderate quality, and two were rated as low. Of the four high-quality studies, one was an RCT **(Type A)**, and the other three were QEDs **(Type B).** All of the eight moderate quality studies fell within **Type C impact evaluations**, and the two low quality studies **fell within Type D impact evaluations,** resulting in an initial evidence security rating of **Level 3**.

While substantial heterogeneity is evident in the meta-analysis, reflected in the RVE adjustment and broad range of observed effects, subsequent moderator analyses account for much of this variation. As a result, the initial evidence

security rating was not downgraded, and **an evidence security rating of 3 out of 5 is maintained.**

All outcomes

Our confidence in the findings on crime-related outcomes is moderate. The meta-analysis included 81 outcomes across 16 studies assessing the broader impact of POP on CYP. These outcomes spanned various domains including offending, victimisation, wellbeing, education, and family relationships.

Study quality varied, with five studies rated as high quality, nine studies rated as moderate quality and two rated as low based on the YEF-EQA. The high-quality studies included one RCT (**Type A**) and four QEDs (**Type B**). All of the moderate quality studies fell into **Type C** (one RCT, six QEDs, one OQD, and one PPD); and the two low quality studies (both PPDs) were designated to **Type D** these combined designs contributed to an initial evidence security classification of **Level 3.**

Similar to the violence analysis, heterogeneity was very high (I^2 = 99.4%), with significant variation across outcomes and study contexts. While this could be seen as a potential reason to downgrade the evidence, the findings from extensive moderator analyses (including five meta-regression models) helped explain a large proportion of this variation particularly through differences in study design, outcome type, setting, and target population. Sensitivity analyses also confirmed the stability of the overall findings, and no individual study or outcome unduly influenced the results.

Therefore, given the general moderate – high evidence base[19], the moderate number of studies, presence of two RCTs, extensive analytical adjustments, and explained heterogeneity, the evidence is not downgraded. We, therefore, **maintain an evidence security rating of 3 out of 5 for the full body of crime-related outcomes.**

The effectiveness of POP interventions varies significantly across the moderators tested, and much statistical heterogeneity was explained by model 4 -

[19] The presence of just one additional type A or B study would have led to an initial evidence security rating of 4.

suggesting that their impact is highly dependent on local context, the specific components implemented, and the fidelity of execution. Understanding what drives successful implementation is crucial for maximising the benefits of POP. The next section explores the key factors that influence implementation, including barriers, facilitators, and real-world challenges in delivering these interventions effectively.

Who does POP work for?

Five of seventeen included studies provided detail in relation to personal characteristics which help us to understand who POP works for (Beito & Sigler, 1997; Braga, 2008b; Braga et al., 2008; Mazerolle et al., 1998; Stokes et al., 1996). These covered ethnicity and experience of deprivation. All five studies were from the US. One of these studies were rated as high quality (Mazerolle et al., 1998); and four as moderate quality (Beito & Sigler, 1997; Braga, 2008b; Braga et al., 2008; Stokes et al., 1996) on the YEF-EQA. Studies where personal characteristics of the sample were described (e.g. gender), but these studies do not develop our understanding of who POP works for, have not been included in this section. No studies explored SEND, care-experience or neurodiversity. No sub-group analyses were possible with the available data.

Ethnicity

Three studies, all from the US, reported on the ethnicity of the population receiving the intervention (Braga, 2008b; Braga et al., 2008; Stokes et al., 1996).

Braga et al. (2008) utilised POP in Lowell, Massachusetts to target gang-involved young people. According to the 2000 U.S. Census, Lowell had 105,167 residents that were 68.6% White, 16.5% Asian, 4.2% Black, and 10.7% mixed or other race (of which, 14% of Lowell residents considered themselves Hispanic). According to Braga et al., there were 19 active street gangs, defined as "a self-identified group of youth who act corporately (at least sometimes) and violently (at least sometimes)". Hispanic and Asian young people dominated Lowell's gang scene. Braga et al. first reported on the initiation of POP in Lowell in 2006, highlighting that the working group were very confident about intervening with Hispanic gangs, but less so with Asian gangs in Lowell, due to their tendency to be more organised, secretive, less territorial and visible. To overcome this, POP targeting Asian gangs focused on gambling businesses run by older Asian gang members, varying from serving search warrants on businesses housing the illegal enterprises to placing a patrol car in front of suspected gambling locations. However, it is important to note that this study was conducted in the US where gangs are more likely to be divided

according to ethnicity than in the UK, questioning the relevance of adapting POP according to group ethnicity in the UK.

Braga (2008b) examined the use of POP in Stockton, California, to respond to gun violence among gang-involved offenders. According to the 2000 U.S. Census, Stockton was the thirteenth largest city in California with 243,771 residents that were 43% White, 20% Asian, 11% Black, and 24% mixed or other races. Nearly one-third of Stockton residents self-identified as Hispanic or Latino origin. Like Braga et al.'s (2008) research in Lowell, Massachusetts, gangs in Stockton comprised of specific ethnic groups: Asian gangs, Hispanic gangs, and African American gangs. The use of POP was deemed successful with a 42% decrease in the monthly number of gun homicide incidents. Unlike the research in Lowell, Braga (2008b) did not describe needing to use any different POP approaches based on the ethnic composition of the gangs. Critically, it is challenging to make comparisons across studies as Braga (2008b) did not provide a definition of a gang.

Stokes et al. (1996) explored the use of POP to support safe travel to and from school in Philadelphia, an area affected by daily incidents of school-related violence. The population of the 'test' school included 846 students during the 1990-1991 school year, with 90% of the student body being African American and the remaining 10% comprised of Hispanic students. Findings suggest no significant change in victimisation or fear of being attacked, questioning the effectiveness of POP with a predominantly African American population. However, only 27.4% of students knew about the 'Safe Corridor' (i.e., the area patrolled by police to enable CYP to go to and from school safely) and only 25.3% used the corridor, which may have reduced its effectiveness.

Whilst these studies provided details on the ethnicity of populations targeted, POP uses a place-based approach and tends to target whole communities. It is *likely* that in other studies included in the effectiveness assessment that POP targeted areas where people from various ethnicities live and congregate.

Experience of Deprivation

Whilst the experience of deprivation was not explicitly communicated in all included studies, most referenced basing POP in urban areas with high rates of crime, with the communities exposed to this approach also likely to have experienced some form of deprivation. Five studies explicitly referred to deprivation, all of which were conducted in the United States (Beito & Sigler, 1997; Braga, 2008b; Braga et al., 2008; Mazerolle et al., 1998; Stokes et al., 1996) but none provided an analysis linking the experience of deprivation to the effectiveness of POP. No research in the UK has explicitly considered the use of POP in areas with increased deprivation.

To summarise, little is known regarding who POP works well for. The evidence base is limited, with a lack of research exploring the characteristics of populations receiving POP, which could help us to understand who POP works for. The little research that has been conducted thus far is all based in the US and relies on descriptive statistics, rather than conducting sub-group analyses.

What factors affect implementation?

Fourteen studies provided evidence related to implementation, of which ten also provided effectiveness data used in the meta-analysis above (see Appendix 5 for details of the studies providing evidence on implementation). Five studies were from the UK (Andrews, 2024; Bullock et al., 2023; Lancashire Constabulary, 2008; Preston Early Intervention Partnership, 2012; Transport for London, 2008), with the remainder from the United States. Two studies were classed as very low quality (Andrews, 2024; Glover, 2002), two were appraised as low (Gulf Breeze Police Department, 2004; Transport for London, 2008) and six as moderate quality (Beito & Sigler, 1997; Braga et al., 2008; Bullock et al., 2023; Maguire et al., 2003; Preston Early Intervention Partnerships, 2012; Weisburd et al., 2008). The four remaining studies included in the implementation section were classed as high quality (Lancashire Constabulary, 2008; Mazerolle et al., 1998; Taylor et al., 2024).

Factors that influenced POP implementation are organised using Proctor's Implementation Outcome Framework (2011). Overall, evaluations of implementation outcomes focused more on the factors that hindered successful implementation with information on factors that support POP implementation lacking. Appendix 6 highlights the availability of evidence according to each of Proctor's implementation outcomes. Where studies reported on the experiences or perspectives of CYP, parents/carers, and professionals, these views are summarised with appropriate direct quotations from primary studies given where available.

To briefly summarise, key themes from this section highlight that for POP to be most effective and accepted within the community, the following should be established during implementation:

- Ensure all decisions made are evidence-led, drawing on appropriate formal analyses of target areas. The use of crime mapping software and drawing on the advanced skills of analysts should be considered. Informal and superficial analyses based on police officers' personal experiences in the target area should be avoided.

- Be flexible and willing to adapt interventions according to the needs of different communities and locations. POP responses which work in one community or area may not be appropriate for others, meaning POP should not be implemented in overly large geographic areas.
- Any responses/interventions implemented should be based on best practice guidelines and research evidence. Programmes that are known to be harmful to intervention participants, such as "Scared straight" strategies, should not be used.
- Consider the concerns of community members, drawing on a collaborative approach to plan POP responses and interventions.
- Close collaborations between police and trusted community stakeholders, such as schools and local faith organisations, can help facilitate engagement in POP responses/interventions.
- Work with key organisations that could be affected by increased referrals from police, such as social services, at an early stage can support with ensuring capacity is available to handle these.
- Support police officers and managers to understand the importance of a POP approach, establishing the benefits of this beyond enforcement alone.
- Monitor that those responsible for implementing POP are fully and consistently adhering to the approach.
- Embed evaluation of the POP approach on an ongoing basis.

Acceptability

Andrews (2024) highlighted that drawing on concerns from community members to decide on locations for interventions, and visibly tackling these, supported their acceptance in the community. Indeed, ensuring all decisions were intelligence-led avoided any one demographic being targeted, increasing POP's acceptability within the community. Supporting this, Maguire et al. (2003) found that using a collaborative approach for intervention planning, led to positive relationships between police, teachers, and young people and increased acceptability of POP. Similarly, Transport for London (2008) indicated that both public perceptions and media discourse on increased crime attributed to young people's widened

access to free bus transport may have enhanced the acceptability of focused efforts to tackle the problem.

Beito and Sigler (1997) highlighted that POP was perceived as an acceptable approach to community members. All community members referred for additional support by the officers (e.g., to social services) made contact with the recommended agencies. Further supporting the acceptability of POP, social services agencies recommended the expansion of the programme given the increase in clients they were able to access and serve from officers' referrals.

The intervention context can impact on the acceptability of a POP approach. For instance, Glover (2002) highlighted that this can be problematic in schools, where education and police departments are perceived as incompatible organisations by both members of each organisation and the wider public. Schools and police can approach negative behaviours in different ways (as a disciplinary vs. criminal issue), impacting on potential collaborations and responses to behaviour. This was found to be problematic in Maguire et al.'s (2003) study examining the implementation of POP in a US school. However, when implemented well, teachers can find POP implemented in schools to be an acceptable approach, as demonstrated by positive teacher feedback following police-led seminars aimed at improving knowledge on drug use among young people in the community (Gulf Breeze Police Department, 2004).

In one UK study (Bullock et al., 2023), findings from interviews with 44 professionals who had experience of implementing POP to target knife crime, indicate that commitment to a problem-oriented approach had not permeated throughout all police or police forces. In general, participants were supportive of using POP but felt that others perceived that enforcement often remained the preferred approach. Engaging police in prevention strategies can be challenging as *"there's still that perception… from some people that we can arrest our way out of serious violence and knife crime. But we can't, and ultimately, we need that different approach to it"* (page 7).

Adoption

Evidence from the UK highlights that having intelligence on areas and times associated with violence involving young people/offending supported the adoption of POP (Bullock et al., 2023), whilst challenges in gaining useful intelligence quickly enough acted as a barrier to implementing POP (Transport for London, 2008). Inconsistent support and scepticism from staff responsible for implementing parts of a POP initiative (including schools, teachers, and transport staff) also hindered implementation within the UK (Transport for London, 2008). Lancashire Constabulary (2008) experienced some time-slippage in the planned implementation and adoption of POP, particularly regarding implementing situational crime prevention responses (i.e., making physical changes in the area) and court processing of enforcements. In addition, changes in staffing impacted upon the continued adoption of POP temporarily, whilst additional interventions were also added that had not been agreed as part of the collaborative approach between police, students and teachers (Maguire et al., 2003).

Being flexible and willing to adapt the interventions according to the needs of different communities and locations, was identified as a key facilitator for the adoption of POP in the UK (Bullock et al., 2023). This was consistent with findings from the US, where interventions were tailored to meet the needs of particular gangs (as outlined in the ethnicity section above) and the extent to which young people were embedded within them (Braga et al., 2008). For example, most violence was reduced through heightened levels of local police patrols and probation, but some heavily embedded members needed enhanced enforcement from federal authorities.

Close collaborations between partners responsible for implementing POP (e.g., schools and police), effective implementation planning, clear and timely feedback between partners, managerial support for the programme, information management, and use of evidence-based approaches were all flagged as facilitators for the adoption of POP (Glover, 2002). Furthermore, Mazerolle et al. (1998) found that continued engagement of partners within problem-solving teams supported the adoption of POP approaches.

The Gulf Breeze Police Department (2004) found that support from the media and local faith organisations to publicise drug and alcohol problems in the community encouraged uptake of seminars/training. In addition, active and engaged staff supporting the POP approach strengthened its adoption within the community. A key barrier to adoption may be the perception among community members that there is no problem to be addressed, for example the Gulf Breeze Police Department (2004) found difficulties in engaging people in the intervention as many did not perceive teenage drug and alcohol use to be a problem in the community.

Strained relationships between community members (including young people) and police (Taylor et al., 2024) can also negatively impact on the adoption of a POP approach. However, this was based on evidence from the US, where relationships between police and community were affected by both the George Floyd murder and a lack of police presence during COVID-19.

Appropriateness

In a UK study of professionals with experience implementing POP, the importance of context was emphasised. Specifically, what works in one context may not work in others: *"you have got to remember the context, because what works in London won't work in Burnley, do you know what I mean? I think we get obsessed with upscaling … and finding that golden solution, but I don't think there is one [for knife crime]".* (page 7, Bullock et al. 2023).

Evidence from the UK suggests that deployment of interventions can be based on superficial analysis of different command areas, with more advanced analytic skills needed to ensure appropriate interventions are used to target specific areas (Andrews, 2024). Despite this, using a POP approach to tackle knife crime ensured that any stop-and-search strategies were appropriately used, with the rate of searches resulting in finds consistently higher than the national average (Andrews, 2024). The need for staff to be adept at using crime mapping software was also reflected in evidence from the US, with this enabling appropriate levels of intervention to be implemented according to the degree of criminal behaviour in

different areas (Braga & Schnell, 2013). Some staff rely on informal assessments of crime levels based on their personal experiences with an area (as opposed to crime mapping tools), which impacts upon the appropriate use of POP (Braga & Schnell, 2013).

There need to be clear records of how officers made decisions within the intervention, including who/where needed targeting and how, to ensure an intervention is appropriate. However, this was lacking in some research (e.g., Beito & Sigler, 1997), whilst other authors highlighted that some taskforces lacked confidence in using POP to target specific groups appropriately (e.g., Braga et al., 2008).

Critically, deploying POP in overly large geographic areas prevents targeted interventions, resulting in police using a 'one size fits all' strategy, which is not consistent with or appropriate for a POP approach (Weisburd et al., 2008). Glover (2002) suggested that schools can be appropriate areas to target for POP interventions, as they have a defined geographical location, relatively stable membership, clear operating hours, clear social structures and roles, defined resources, high visibility in the community and acceptance as a legitimate institution.

To support community members to recognise the appropriateness of a POP approach, Gulf Breeze Police Department (2004) utilised a collaborative approach to develop an action plan relevant to the local area. Law enforcement, school administrators, active citizens, media, teachers and student representatives helping all aided in the creation of the action plan.

Feasibility

Bullock et al. (2023) highlighted several factors negatively affecting the feasibility of implementing POP in the UK. This included difficulties analysing crime data at speed due to limited availability of analysts, poor quality of available data, reliance on limited and potentially biased police data, and difficulties accessing data from other organisations (e.g., emergency health services). In addition, Transport for London (2008) highlighted that the high numbers of young people

using their service (accounting for 1,100,000 journeys a day on the bus network) impacted upon the feasibility of rolling this out across the network to target all potential users. As such, a gradual rollout was planned, enabling the intervention to be modified as it progressed.

Glover (2002) suggested that it is important to differentiate between community- and school-based interventions for it to be feasible to implement POP in schools. Specifically, school administrators considering implementing POP in schools need clarity on the extent to which any approach extends beyond the boundaries of the school and the school environment. Gulf Breeze Police Department (2004) did implement a POP approach in schools, finding that it was a feasible approach, with only 64 hours of overtime needed across the school year. As a result, the collaborative approach was planned to continue with new classes each year.

Fidelity

Police personnel in the UK, interviewed by Bullock et al. (2023), highlighted that it is difficult to assess the fidelity of POP approaches due to a lack of evaluations of implemented programmes. There was a lack of fidelity of in Transport for London's (2008) POP approach, with resistance from employees and schools meaning the approach was inconsistently applied.

Two studies from the US reported that the 'dosage' (degree of contact between community and police) may not have been high enough to lead to positive outcomes (Taylor et al., 2024; Weisburd et al., 2008). Indeed, Taylor et al. (2024) flagged that 16-17 visits to hotspot areas per month to implement POP was recommended by police leadership, but only 1.3-7.4 visits were conducted per hotspot area. A further study from the US suggested that the number of interventions implemented varied per high crime area, dependent upon leaders' commitment to a POP approach (Braga & Schnell, 2013). For low-fidelity areas, command staff had to be provided with additional training and closer supervision.

Reach/Penetration

Police personnel from the UK highlighted that most police forces across England and Wales have, at some point, used POP (Bullock et al., 2023). However, Bullock et al. (2023) found that POP is inconsistently applied across the country.

Critically, there was a lack of information available across the included studies on reach/penetration.

Sustainability

Bullock et al.'s (2023) research with UK professionals indicated that the sustainability of POP was challenging, with continued commitment to deliver this approach affected by its perceived high cost and risks associated with losing fundings.

Beito and Sigler's (1997) research highlighted that continued use of POP affects the capacity of other agencies too. For example, with increased referrals to social services from officers, there would need to be additional resources available for these services, should a POP approach continue.

Gulf Breeze Police Department (2004) reported that it would take several years of implementing a POP approach to control drug and alcohol use with each new school group targeted separately, but that it remained a sustainable approach because of its positive effects. They highlighted that for it to remain a sustainable approach, it would have to be embedded within annual operations of both school and police departments.

Experiences of children and young people

Only one study reported on the experiences of children and young people. Preston Early Intervention Partnership (2012) found that 45% of CYP said they felt 'scared' or 'terrible' following a visit to a custody suite as part of the POP approach, with a further 31% feeling 'cold and lonely', 'worried' or 'sad'. However, this aspect of the wider POP approach was inspired by 'scared straight', which was unique to the

Preston Early Intervention POP approach. This scared straight approach has been widely discredited in the literature.

How much does it cost?

Within this systematic review, no authors provided details on the actual costs incurred.

One UK paper (Preston Early Intervention Partnership, 2012) identified a potential cost saving of £82,000, due to a reduction in arrests. Each arrest was estimated to cost £2000, with a reduction in arrests equating to £130,000 in 2008/9 to £48,000 in 2010/11. An additional paper from a UK police force (Lancashire Constabulary, 2008) identified considerable cost savings after implementing POP, including $62,100[20] associated with a reduction in burglary, $72,420 due to a reduction in criminal damage, and $51,770 related to a reduction in antisocial behaviour.

One paper from a US police department (Gulf Breeze Police Department, 2004) highlighted that the core costs incurred were the project personnel and time. Most activities occurred during normal scheduled work time and did not incur additional costs, apart from approximately 64 hours of overtime across the year-long initiative (covering evening bicycle patrols, canine officers and parental seminars). Given the effectiveness of the program at reducing crime, Gulf Breeze Police Department felt the costs were justifiable.

[20] Please note, despite this being based in the UK, cost savings were reported in dollars.

Conclusion and Takeaway Messages

There is strong evidence from this systematic review and meta-analysis to suggest that Problem-Oriented Policing (POP) is effective in reducing violence among children and young people and in improving their lives across a range of additional outcomes.

POP interventions resulted in an estimated -**23.9% reduction in violent crime** outcomes, highlighting its potential as a targeted approach to addressing violence involving young people. These findings are based on a meta-analysis of 11 studies and 38 outcomes. The broader meta-analysis of 16 studies and 81 crime-related outcomes found a Relative Incident Rate Ratio (RIRR) of 0.80 (95% CI 0.74 – 0.84), which corresponds to a **21.0% reduction in crime and associated negative outcomes**.

These results are broadly consistent with previous systematic reviews on general populations, such as those conducted by Hinkle et al. (2020, 2024), which found that POP was most effective for property crime (31% reduction) and disorder (18.9% reduction). In the Hinkle analysis, the effects on violent crime were smaller (9.5% reduction) and not statistically significant in the main analysis. However, in an alternative analysis where they included the largest effect size from studies with multiple outcomes, a significant 18.3% reduction in violent crime was reported.

A recent meta-analysis from Turchan and Braga (2024), showed that Problem-Oriented Policing in crime hot spots led to a 24% reduction in violent crime in treated places compared to control areas, which is closely aligned with our findings (23.9%). A further breakdown of the type of violent crime suggested varying effectiveness:

- Robbery: 20% reduction
- Assault Crimes: 29% reduction
- Firearm-Related Violent Crimes: 36% reduction

The confidence in findings is moderate, made stronger still by the clear alignment with other, larger meta-analyses conducted on the general population. Two

Randomised Controlled Trials (RCTs) provided robust causal evidence, but most studies relied on quasi-experimental designs and other quantitative methods, which limit definitive conclusions about causality. Despite these limitations in the evidence base itself, the robust systematic meta-analysis provides clear evidence that POP can significantly reduce violence and crime among children and young people. However, its effectiveness likely varies based on context, intervention design, and fidelity of implementation.

What works?

Our analysis highlights key factors that influence the effectiveness of POP:

- Moderator analysis #5 demonstrates that POP interventions were more strongly associated with improvements in personal development and wellbeing, while direct crime and offending outcomes showed weaker, non-significant effects.
- Targeted interventions aimed at gang-involved young people were more effective than broader area-based interventions. This aligns with findings from Hinkle et al. (2024), which emphasise that high-risk group targeting yields stronger effects.
- Fidelity is critical—when interventions were scaled up across a wider area, fidelity was often compromised, leading to a shift toward traditional enforcement tactics rather than a problem-solving approach. This suggests that without structured implementation support, POP interventions may not achieve their full potential.

Who Benefits Most?

Despite its general effectiveness, there is a lack of high-quality evidence on who benefits most from POP interventions, for example, no research specifically examined its impact on neurodiverse young people, young people with care experience, or those with SEND. This is a major gap in the evidence base and highlights the need for data-driven targeting (using clear analytic approaches)

rather than reliance on informal assessments, which risk reinforcing systemic biases.

Our meta-analysis found that focused interventions targeting young people involved in gangs were more effective than broader interventions aimed at the general population. However, the studies included in our review do not provide clear explanations for this trend. Previous meta-analyses by Hinkle et al. (2024) and Weisburd et al. (2010) suggest that POP is most effective when interventions are precisely tailored to specific populations rather than being applied broadly. This highlights the importance of developing targeted, evidence-based strategies to maximise the impact of POP interventions.

Limitations

Challenges with the studies measuring effectiveness

This report highlights several key challenges in evaluating the impact of Problem-Oriented Policing, including heterogeneity in study designs, difficulties in standardising effect sizes, and complexities in measuring intervention effectiveness. A major issue was the lack of standardised effect sizes in the original studies, often requiring manual calculations from reported crime counts. Additionally, adjusting for over-dispersion in count data further complicated effect size estimation, while variations in crime definitions and intervention methodologies made direct comparisons difficult.

Importantly, the application of Robust Variance Estimation (RVE) led to changes in p-values and statistical significance for some outcomes, reflecting increased uncertainty due to intra-study dependence. This highlights the need for cautious interpretation and supports the robustness of findings that remained significant after RVE adjustment.

These challenges are not unique to this analysis. Similar issues have been identified in other meta-analyses, including Hinkle et al. (2020, 2024), Turchan & Braga (2024), and Weisburd et al., (2010), underscoring a broader problem in policing research. Inconsistent reporting and methodological limitations continue

to hinder clear conclusions on intervention effectiveness and reduce the comparability of findings across studies. Addressing these limitations through more standardised and transparent reporting practices is essential for improving the reliability of future meta-analyses in crime and policing research.

Challenges with the studies measuring implementation

This report highlights that there is a lack of detailed reporting on implementation processes across studies. In particular, data regarding the feasibility, fidelity, reach and penetration, and sustainability of POP was lacking. For programmes that did report on fidelity, it was clear that there was resistance from staff and inconsistent implementation of a POP approach. However, this was only discussed briefly in three of the 11 included papers, meaning it is difficult to assess the impact of programme fidelity on outcomes. In addition, challenges include the lack of detail on how decisions are made on where to implement POP approaches, how specific interventions are decided upon, and the degree of stakeholder and community engagement in devising and implementing a POP approach.

This is consistent with previous meta-analyses which highlighted that some officers viewed POP as outside their traditional role, leading to weak implementation (Hinkle et al., 2020), while resistance and inconsistent commitment from leadership delayed or derailed interventions (Weisburd et al., 2010). Similar to our findings, both of these past reviews highlight that these issues were often mentioned only in passing, rather than systematically assessed, making it difficult to quantify their impact on effectiveness.

Final thoughts and Recommendations

Problem-Oriented Policing (POP) offers a promising strategy for reducing violence involving children and young people in the UK, but its successful implementation and evaluation remain challenging. The success of the approach is likely to rely on bespoke, multi-agency interventions involving specially trained staff, targeting deep-rooted social and environmental factors that contribute to crime. These complexities may present barriers to adoption, particularly in resource-limited

settings. To fully realise the potential of POP, sustained investment in high-quality research, data-sharing mechanisms, and police training is essential. Based on our understanding of previous meta-analyses and this current research, we make the following key recommendations:

1. More **high-quality evaluations** of POP are needed to determine what works best and for whom. As Hinkle et al. (2024) emphasise, POP evaluations must document interventions in detail, rather than treating them as a "black box."

2. We noted in our data extraction, too many interventions base their data collection on informal assessments, likely reinforcing systemic biases rather than targeting actual crime drivers. We suggest that personnel who implement POP improve on, and create, **robust data-sharing** between police, social services, and education providers. This recommendation is supported by Turchan and Braga (2024) who stress that standardised data collection frameworks are crucial to improving intervention effectiveness and comparability across studies.

3. Our review found some strong evidence that targeted interventions (e.g., for gang-involved young people) were more effective than broad, unfocused approaches. This aligns with Weisburd et al. (2010), who found that narrowly defined, data-driven interventions yield better results. To fully realise the impact of POP in reducing violence involving young people, we recommend **future investment in targeted, evidence-based interventions** that focus on high-risk groups rather than broad, general population approaches.

References

Studies marked with an asterisk (*) were included in the effectiveness and implementation assessments.

*Andrews, T. (2024). Deter, detect, disrupt. An analysis of Nottinghamshire Police's knife crime team. The Police Journal: Theory, Practice and Principles. https://doi.org/10.1177/0032258x241286359

*Baker, T., & Wolfer, L. (2003). The crime triangle: Alcohol, drug use, and vandalism. Police Practice and Research, 4(1), 47–61. https://doi.org/10.1080/1561426032000059187

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. Biometrics, 50(4), 1088–1101. https://doi.org/10.2307/2533446

*Beito, L. R., & Sigler, R. T. (1997). Police community service: An evaluation of program effectiveness. International Journal of Comparative and Applied Criminal Justice, 21(1), 151–162. https://doi.org/10.1080/01924036.1997.9678592

*Boston Police Department. (2008). District D-14: Breaking and entering solution plan. https://popcenter.asu.edu/sites/default/files/library/awards/goldstein/2008/08-03(F).pdf

*Braga, A. A., McDevitt, J., & Pierce, G. L. (2006). Understanding and Preventing Gang Violence: Problem Analysis and Response Development in Lowell, Massachusetts. Police Quarterly, 9(1), 20–46. https://doi.org/10.1177/1098611104264497

Braga, A. A. (2008a). Problem-Oriented Policing and Crime Prevention (2nd ed.). Criminal Justice Press.

*Braga, A. A. (2008b). Pulling levers focused deterrence strategies and the prevention of gun homicide. Journal of Criminal Justice, 36(4), 332–343. https://doi.org/10.1016/j.jcrimjus.2008.06.009

*Braga, A. A., Kennedy, D. M., Waring, E. J., & Piehl, A. M. (2001). Problem-Oriented Policing, Deterrence, and Youth Violence: An Evaluation of Boston's Operation Ceasefire. Gangs, 513–543. https://doi.org/10.4324/9781351157803-28

*Braga, A. A., Pierce, G. L., McDevitt, J., Bond, B. J., & Cronin, S. (2008). The Strategic Prevention of Gun Violence Among Gang-Involved Offenders. Justice Quarterly, 25(1), 132–162. https://doi.org/10.1080/07418820801954613

*Braga, A. A., & Schnell, C. (2013). Evaluating Place-Based Policing Strategies. Police Quarterly, 16(3), 339–357. https://doi.org/10.1177/1098611113497046

Braga, A. A., Welsh, B. C., & Schnell, C. (2019). Disorder policing to reduce crime: A systematic review. *Campbell Systematic Reviews*, *15*(3), e1050.

*Bullock, K., Agar, I., Ashby, M., Brennan, I., Hales, G., Sidebottom, A., & Tilley, N. (2023). Police practitioner views on the challenges of analysing and responding to knife crime. CrimRxiv. https://doi.org/10.21428/cb6ab371.2eeed9c4

Cochran, W. G. (1954). The combination of estimates from different experiments. Biometrics, 10(1), 101–129. https://doi.org/10.2307/3001666

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement,* 20(1), 37–46. https://doi.org/10.1177/001316446002000104

College of Policing (2020). Problem-oriented policing. https://www.college.police.uk/research/crime-reduction-toolkit/problem-oriented-policing [Accessed 20/03/2025]

Eck, J. E. & Spelman, W. (1987). Problem Solving: Problem-Oriented Policing in Newport News. Police Executive Research Forum. https://www.ojp.gov/pdffiles1/Digitization/111964NCJRS.pdf

*Glover, R. (2002). Community and Problem Oriented Policing in School Settings: Design and Process Issues.

*Gulf Breeze Police Department. (2004). Excellence in Problem-Oriented Policing Award Submission. https://popcenter.asu.edu/sites/default/files/library/awards/goldstein/2004/04-13.pdf

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39-65. https://doi.org/10.1002/jrsm.5

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21(11), 1539–1558. https://doi.org/10.1002/sim.1186

Hinkle JC, Weisburd D, Telep CW, & Petersen K. (2020). Problem-oriented policing for reducing crime and disorder: An updated systematic review and meta-analysis. *Campbell Syst Rev*. 16:e1089. https://doi.org/10.1002/cl2.1089

Hinkle JC, Weisburd D, Telep CW, & Petersen K. (2024). When is problem-oriented policing most effective? A systematic examination of heterogeneity in effect sizes for reducing crime and disorder, *Policing: A Journal of Policy and Practice*, Volume 18, paae053, https://doi.org/10.1093/police/paae053

*Lancashire Constabulary. (2008). *'MOPPIN up Dodge'.* Goldstein Application.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics,* 33(1), 159–174. https://doi.org/10.2307/2529310

*Maguire, E. R., Schnobrich-Davis, J., Solomon, S. E., & Uchida, C. D. (2003). Reducing Bullying, Threats, and Intimidation in Westwood, MA: An Evaluation of a Problem Solving Partnership. https://cops.usdoj.gov/pdf/school_based/Westwood_MA.pdf

*Mazerolle, L. G., Ready, J., Terrill, W., & Gajewski, F. (1998). Problem-oriented policing in public housing: Final report of the Jersey City Project.

https://www.nhlp.org/wp-content/uploads/Mazerolle-Problem-Oriented-Policing-in-Public-Housing-1998.pdf

Modise, J. M. (2023). Violence prevention, general deterrence, community policing, and problem-oriented policing are all components of hot spot policing. *International Journal of Innovative Science and Research Technology*, 8(4), 2619-2635. https://www.researchgate.net/publication/376260593

*Preston Early Intervention Partnership. (2012). *The Custody Experience: Reducing 1st time entrants into the Criminal Justice System.* Herman Goldstein Awards.

Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health and Ment Health Serv Res.* 38:65–76. doi: 10.1007/s10488-010-0319-7

Pustejovsky, J. E. (2024). *clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections* (Version 0.5.11) [R package]. https://CRAN.R-project.org/package=clubSandwich

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. Prevention science : the official journal of the Society for Prevention Research, 23(3), 425–438. https://doi.org/10.1007/s11121-021-01246-3

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. British Medical Journal, 342, d549. https://doi.org/10.1136/bmj.d549

Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), Publication bias in meta-analysis: Prevention, assessment and adjustment (pp. 99–110). Wiley.

*Stokes, R., Donahue, N., Caron, D., & Greene, J. R. (1996). Safe Travel to and From School: A Problem- Oriented Policing Approach [dataset]. In PsycEXTRA Dataset. American Psychological Association (APA). https://doi.org/10.1037/e519862006-001

*Taylor, B. G., Liu, W., Zalsha, S., Sheridan-Johnson, J., Sabol, G., & Lacy, C. R. (2024). The impact of a youth-focused problem-oriented policing initiative on crime: findings from a randomized controlled trial in three cities. Journal of Experimental Criminology. https://doi.org/10.1007/s11292-024-09629-6

Telep, C. W., & Hibdon, J. (2019). Understanding and responding to crime and disorder hot spots (Problem-Solving Tool Guide No. 14). ASU Center for Problem-Oriented Policing. https://popcenter.asu.edu/content/tool-guides-crime-and-disorder-hot-spots

Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics*, *40*(6), 604–634. https://doi.org/10.3102/1076998615606099

*Transport for London. (2008). A problem-oriented policing approach to tackling youth crime and anti-social behavior on London's buses. https://popcenter.asu.edu/sites/default/files/library/awards/goldstein/2008/08-47(F).pdf

Turchan B and Braga AA. (2024). The effects of hot spots policing on violence: A systematic review and meta-analysis, *Aggression and Violent Behavior*, https://doi.org/10.1016/j.avb.2024.102011

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. Journal of Educational and Behavioral Statistics, 30(3), 261–293. https://doi.org/10.3102/10769986030003261

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. Journal of Statistical Software, 36(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. Research Synthesis Methods, 1(2), 112–125. https://doi.org/10.1002/jrsm.11

*Weisburd, D., Morris, N. A., & Ready, J. (2008). Risk-Focused Policing at Places: An Experimental Evaluation. Justice Quarterly, 25(1), 163–200. https://doi.org/10.1080/07418820801954647

Weisburd, D., Telep, C.W., Hinkle, J.C. and Eck, J.E. (2010). Is problem-oriented policing effective in reducing crime and disorder?. Criminology & Public Policy, 9: 139-172. https://doi.org/10.1111/j.1745-9133.2010.00617.x

Wilson, D. B. (2022). The relative incident rate ratio effect size for count-based impact evaluations: When an odds ratio is not an odds ratio. Journal of Quantitative Criminology, 38(2), 323–341.

Appendix 1. Methods of this systematic review

Protocol

Prior to initiating this systematic review, we developed a comprehensive protocol for an Evidence and Gap Map (EGM) outlining the research objectives, eligibility criteria, search strategy, data extraction, quality appraisal, and synthesis methods. This protocol was registered and is available on the Open Science Framework (OSF),[21] ensuring transparency and adherence to predefined methods.

The search strategy and eligibility criteria outlined in the protocol are designed to be sufficiently comprehensive to capture a broad and systematically identified body of literature, enabling the extraction of relevant subsets of studies for inclusion in the Toolkit. The methods described below are aligned with the current Toolkit Strand on Problem-Oriented Policing ensuring a structured and rigorous approach to evidence synthesis.

Details of screening and Interrater reliability

A total of 1,956 titles and abstracts identified as potentially relevant to the current strand were independently assessed by two reviewers. To ensure a fair distribution of workload, the screening process was structured as follows:

- Three reviewers each screened 500 records.
- One reviewer screened 456 records.
- The project manager (Reviewer 1) conducted a duplicate screening of all 1,956 records to enhance consistency.

To evaluate the consistency of screening decisions among reviewers, Cohen's Kappa (κ) (Cohen, 1960) was calculated. The results were interpreted based on Landis and Koch (1977) benchmarks:

[21] Protocol is available to access here: https://osf.io/vamxy

Table A1 11 *Cohen's Kappa and interpretation for reviewer pairs screening by title and abstract*

Reviewer Pair	Cohen's Kappa (κ)	Interpretation (Landis & Koch, 1977)
Reviewer 1 vs. Reviewer 2 *(n=465)*	0.536	Moderate agreement
Reviewer 1 vs. Reviewer 3 *(n=500)*	0.236	Fair agreement
Reviewer 1 vs. Reviewer 4 *(n=500)*	0.056	Slight agreement
Reviewer 1 vs. Reviewer 5 *(n=500)*	-0.456	No agreement

A senior team member with content expertise (Reviewer 2) reconciled discrepancies between Reviewer 1 and Reviewers 3, 4, and 5. Common errors and inconsistencies were noted and discussed in a team meeting, ensuring alignment in decision-making criteria. Additionally, Reviewer 1 and Reviewer 2 conducted a focused review to resolve any outstanding discrepancies between their studies.

At the end of title and abstract screening:

- 117 studies were marked as included.
- 468 studies were marked as maybes.
- 1,371 studies were marked as excluded.

A total of 585 studies proceeded to full-text screening.

Table A2 12 *Full text screening results*

Reason for exclusion	Number of Records Excluded at Full-Text Level
310	Did not target CYP
127	PDF not accessible
58	Study Design not eligible
10	Outcomes or intervention not relevant
2	Excluded based on language
9	Duplicates

For inaccessible PDFs, the team attempted to contact lead authors to request access to the report or further data. Following full-text screening, 69 studies were flagged as potentially relevant for inclusion (after being reviewed by at least two independent reviewers). These were distributed among pairs of reviewers to ensure each paper was extracted and checked by a second person.

Of these 69 papers, 21 papers were excluded upon further review. All 21 excluded studies were thoroughly checked by a senior team member.

Table A3 13 *Reasons for exclusion after full text screening*

Reason for exclusion	Number of Records Excluded at EGM Data Extraction Level
Study design not meeting inclusion criteria	15
Did not target CYP	3
Outcomes or intervention not relevant	3

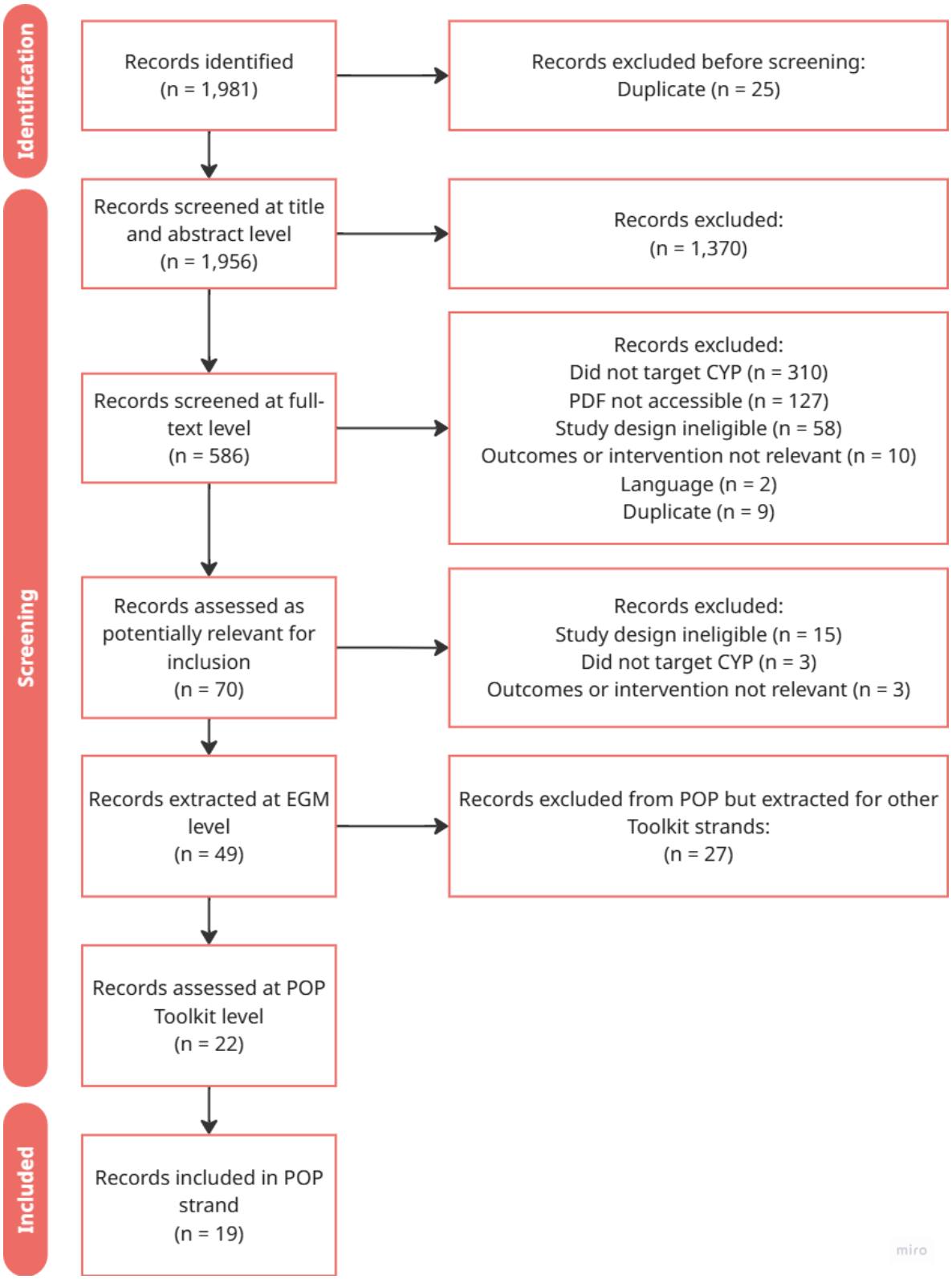
Following data extraction for the EGM (48 studies), 22 studies were deemed eligible for additional extraction for the POP Toolkit strand.

However, of the 22 studies initially selected for further extraction for the POP strand, three were subsequently excluded by a senior member of the research team. One study (Fox, 2015) was allocated to a different research strand on focussed deterrence. Another study (Uchida, 2001) was excluded as it did not target children or young people in the intervention. The third study (Braga et al., 2006) was identified as containing data already included in another report[22] and linked to it via EPPI - Braga et al. (2008).

Therefore, a total of 19 studies were included at the Toolkit data extraction level. The characteristics of these studies are detailed in Appendix 3 and Appendix 5.

[22] In cases where multiple reports originate from the same study, the earlier report is typically categorised as a duplicate in the PRISMA diagram, however, the paper is retained and linked to the report that the review team designates the most comprehensive (or primary) source for inclusion

Figure A1 5 PRISMA flow diagram for the Problem-Oriented Policing strand



Quality appraisal process

The YEF-EQA tool was used across all 19 Toolkit studies to systematically assess the quality, reliability, and relevance of the research. This tool was applied by one reviewer, with a second reviewer checking their appraisals.

Table A4 14 *Quality appraisal ratings for studies included in the Problem-Oriented Policing Toolkit strand*

Study ID	Overall quality of the study	Study Design
Andrews (2024)	Very low	Process Evaluation
Baker & Wolfer (2003)	High	QED – Quasi-Experimental Designs
Beito & Sigler (1997)	Moderate	QED – Quasi-Experimental Designs
Boston Police Department (2008)	Moderate	OQD – Other Quantitative Designs
Braga et al. (2001)	Moderate	QED – Quasi-Experimental Designs
Braga (2008b)	Moderate	QED – Quasi-Experimental Designs
Braga et al. (2008)	Moderate	QED – Quasi-Experimental Designs
Braga & Schnell (2013)	High	QED – Quasi-Experimental Designs
Bullock et al. (2023)	Moderate	Process Evaluation
Glover (2002)	Very low	Process Evaluation

Gulf Breeze Police Department (2004)	Low	PPD – Pre/Post Designs
Lancashire Constabulary (2008)	High	QED – Quasi-Experimental Designs
Maguire et al. (2003)	Moderate	PPD – Pre/Post Designs
Mazerolle et al. (1998)	High	QED – Quasi-Experimental Designs
Preston Early Intervention Partnership (2012)	Moderate	QED – Quasi-Experimental Designs
Stokes et al. (1996)	Moderate	QED – Quasi-Experimental Designs
Taylor et al. (2024)	High	RCT – Randomised Control Trial
Transport for London (2008)	Low	PPD – Pre/Post Designs
Weisburd, Morris, & Ready (2008)	Moderate	RCT– Randomised Control Trial

How the findings were analysed and combined.

Effectiveness data

When working with count data, the relative incident rate ratio (RIRR) provides a robust measure of change over time, independent of variations in population size or time intervals. Wilson (2022) notes that raw counts can be misleading due to population fluctuations, so analysis should focus on changes in underlying rates. A critical property of log RIRR and its logarithmic transformation is the stability across different units of time and population scales (e.g., annual vs. monthly rates or per 100,000 vs. per 10,000 population).

The meta-analysis used the log Relative Incidence Rate Ratio (log RIRR) as the primary effect size metric to assess changes in crime rates over time.

$$RIRR = \frac{(AxD)}{(BxC)}$$

To stabilise variances, the log transformation of RIRR was applied. The variance of log RIRR was computed as:

$$\boldsymbol{V(logRIRR) = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}}$$

To account for over-dispersion in count data, an adjustment was applied:

$$\boldsymbol{V(logRIRR)\ x\ D, \qquad N = 0.0008\ X\ N + 1.2}$$

where N represents the mean number of incidents per case.

A Poisson regression model was used to analyse count data, addressing potential over-dispersion and handling the skewed distribution of crime incidents. The analysis included random effects to account for between-study variability.

Finally, the IRR/RIRR can be converted to a percentage change metric using the following equation:

$$\%\ change = 100\ \times (IRR - 1)$$

Meta-regression

In the current meta-analysis and five meta-regression models, we have chosen to include both single-series (e.g., single-case) and comparative studies to provide a comprehensive synthesis of the available evidence. Recognising that these study designs differ in methodology and interpretation, a moderator analyses has assessed potential discrepancies between the study types. This approach allows the team to identify and account for variations that may arise due to study design differences, ensuring a more nuanced and accurate understanding of the intervention's effects.

Please note that the Meta-regression estimates for models 1-5 do not directly represent log RIRR values but indicate how moderators influence the intervention effect. The intercept reflects the baseline log RIRR, where a negative value

suggests an overall reduction in crime. Moderator estimates show deviations from this baseline, so a positive estimate does not imply an increase in crime but rather a weaker reduction compared to the reference category. For example, an estimate for "Breaking the Law or Offending Behaviour" of 1.25 (p = 0.001) suggests that studies measuring this outcome report effects closer to zero (less effective at reducing crime) than the baseline. Conversely, negative moderator estimates (e.g., Criminal Peers: -0.99, p = 0.035) indicate a stronger reduction in crime, thus more effective, compared to the baseline. It is also important to remember that moderator estimates reflect relative changes in effect size, not absolute increases or decreases in crime.

Preparing the data frame for analysis

As is common in meta-analysis, the team encountered several challenges in harmonising effect sizes across studies. First, there were instances where certain studies reported **zero events** in one or both comparison groups (e.g., zero cases of rape in the control group before intervention; Beito & Sigler, 1997). This posed challenges in calculating effect sizes, as standard methods like the logRIRR become undefined when event counts are zero. To address this, a continuity correction was applied (by simply adding 0.5 to each cell of the 2x2 contingency tables). This approach stabilises variance estimates and allows for the inclusion of studies with zero events, thus maintaining the integrity of the meta-analysis.

Second, there was a need to standardise the direction of log Relative Incident Rate Ratio (log RIRR) values so that negative values consistently indicate a favourable intervention effect[23]. For example, in Baker & Wolfer (2003), the outcome "Feel safe in the park during the night" increased more in the treatment group (from 19.4% to 53.2%) than in the control group (from 33.2% to 58.9%), showing a greater improvement in the treatment condition. Since higher scores reflected a beneficial effect, the log RIRR required a negative adjustment to align

[23] Since a negative log RIRR reflects a decrease in crime (i.e., a reduction in incident rates), we retained the original direction of log RIRR values to align with this interpretation.

with other outcomes where lower incident rates indicated a positive effect[24]. This approach ensures consistent and interpretable results across studies.

Third, there were occasions where results and/ or level of variance was not reported in a way that allowed us to have confidence in the results, in cases where we could make inferences based on other statistics, we included that data in the meta-analysis description but did not attempt to include their data due to potential of bias. Briefly below we highlight those studies:

Gulf Breeze Police Department, 2004 – This Herman Goldstein Award submission reports percentage reductions in offenses but lacks raw incident counts, standard errors, and confidence intervals necessary for calculating precise effect sizes. Consequently, this study was excluded from our meta-analysis. We have reached out to the authors to request additional data.

Mazerolle et al., 1998 - Due to the use of an autoregressive mixed model with extra-dispersion adjustments, the intervention coefficient included (-0.0077, Table 4.1) may not directly represent a log rate ratio. Additionally, we are concerned as the reported 25% reduction in calls to police reported by the author does not align with the available estimates, suggesting that further context or raw incident counts may be needed for accurate calculation. We have reached out to the authors to request additional data for clarification.

Transport for London, 2008 - The document provides percentage reductions in crime and rates per million passenger journeys but does not include raw incident counts pre- and post-intervention, exposure time for calculating rates, or confidence intervals (CIs) and standard errors (SEs) for the reported crime rates. These are necessary for calculating log RIRR and SE for inclusion in a meta-analysis. We have reached out to Transport for London to request additional data.

[24] This is achieved by simply multiplying the positive log RIRR by -1, Standard errors (SE) remained unchanged, as they represent variability rather than effect direction.

Meta-analysis

A random-effects model was fitted to the data. The amount of heterogeneity (i.e., τ^2), was estimated using the restricted maximum-likelihood estimator (Viechtbauer, 2005). In addition to the estimate of τ^2, the Q-test for heterogeneity (Cochran, 1954) and the I^2 statistic (Higgins & Thompson, 2002) are reported. In case any amount of heterogeneity is detected, a prediction interval for the true outcomes is also provided (Riley et al., 2011).

Use of Robust Variance Estimation (RVE)

To account for the statistical dependence that arises when synthesising multiple effect sizes from the same study (e.g., due to multiple outcomes, time points, or subgroups), we employed Robust Variance Estimation (RVE). Standard meta-analytic models assume independence among effect sizes; however, this assumption is often violated in practice, which can lead to underestimated standard errors and inflated Type I error rates (Hedges, Tipton, & Johnson, 2010).

We applied RVE using the clubSandwich package in R (Pustejovsky, 2024), specifically the vcovCR() function with type = "CR2", clustering on Study_ID. This method uses the CR2 small-sample correction, based on the Satterthwaite approximation, which is recommended when the number of clusters is relatively small (Tipton & Pustejovsky, 2015; Pustejovsky & Tipton, 2022). We then used coef_test() to obtain corrected standard errors, confidence intervals, and p-values, which we presented in tables.

We did not use a correlated effects model or specify a correlation matrix for within-study effect sizes. This was due to limitations in available data; detailed within-study correlations were not reported consistently across included studies. While best practice encourages specifying a correlation structure where possible, current methodological guidance also supports the use of study-level clustering with CR2 corrections as a valid alternative when such detail is unavailable (Pustejovsky & Tipton, 2022).

As is typical in meta-analysis, the use of RVE resulted in changes in the statistical significance of some estimates. In the main analysis, several effects that were

statistically significant under a standard REML model became non-significant after applying RVE. This shift indicates that the original model likely underestimated uncertainty due to intra-study dependence. By applying RVE, we provide more conservative and reliable inference. These differences are transparently reported in the results section, with a comparative table outlining where significance levels changed.

Sensitivity analyses

Studentized residuals and Cook's distances are used to examine whether studies may be outliers and/or influential in the context of the model (Viechtbauer & Cheung, 2010). Studies with a studentized residual larger than the $100\times(1-0.05/(2\times k))$th percentile of a standard normal distribution are considered potential outliers (i.e., using a Bonferroni correction with two-sided $\alpha=0.05$ for k studies included in the meta-analysis). Studies with a Cook's distance larger than the median plus six times the interquartile range of the Cook's distances are considered to be influential. The rank correlation test (Begg & Mazumdar, 1994) and the regression test (Sterne & Egger, 2005), using the standard error of the observed outcomes as predictor, are used to check for funnel plot asymmetry. The analysis was carried out using R (version 4.4.2) (R Core Team, 2020) and the metafor package (version 4.8.0) (Viechtbauer, 2010).

Implementation data

Information on factors that influenced, or was perceived to influence, implementation was extracted from studies where this was reported by study authors.

To capture implementation outcomes the toolkit data extraction made use of Proctor et al's (2011) Implementation Outcomes Framework to capture and categorise the barriers and facilitators to achieving good implementation.

The data extraction for the toolkit is an extension of what is already captured in the EGM. For the EGM the focus was on whether or not implementation outcomes

were measured. In other words, does a study report on indicators of how well the programme/intervention was implemented or not. For toolkit data extraction we capture why implementation did or did not go well, what influenced implementation? This is typically thought of as barriers and facilitators to implementation. Information on barriers and facilitators will be presented using Proctor et al's (2011) Implementation Outcomes as headings so that the reader can understand the evidence, and gaps in the evidence, on the following implementation outcomes:

- **Acceptability**: Stakeholders' perceptions that the intervention or change is agreeable, palatable, or satisfactory.
    - Example indicators: Children's views on the intervention, participant engagement, satisfaction with content or delivery.
- **Adoption:** The decision or action to employ an intervention or implementation target.
    - Example indicators: Uptake of the intervention by services, schools, or communities.
- **Appropriateness:** The perceived fit or relevance of the intervention to the given context or problem.
    - Example indicators: Adaptations made to improve the intervention's fit with the context, perceived usefulness.
- **Feasibility**: The extent to which the intervention can be successfully implemented in a specific setting.
    - Example indicators: Evidence of practicality or utility, ability to deliver the intervention in the target environment.
- **Fidelity**: The degree to which the intervention was delivered as intended.
    - Example indicators: Training quality, dosage and intensity of the intervention, adherence to the prescribed approach.

- **Reach/Penetration:** The extent to which the intervention has been integrated into a service setting or reached eligible recipients.
    - Example indicators: Ratio of recipients served to the target population, evidence of saturation or integration.
- **Sustainability:** The ability to maintain or institutionalise the intervention over time.
    - Example indicators: Evidence of routinisation, integration into policies or practices, durability of implementation efforts.

Where implementation barriers/facilitators or influences on an implementation outcome were not measured and/or reported this is stated.

The information extracted on each implementation outcome was narratively summarised. Further analysis and integration of implementation information with the meta-analysis and meta-regression was limited because of a lack of detailed evaluations of implementation.

Appendix 2. Location details

	Number of UK Studies	Number (and Location) of International Studies
Overall, for Strand	3 (one study -TfL excluded on data at the meta-analysis stage)	13 (United States)
Contributing to Evidence Quality Rating	1	12 (United States)
Contributing to Estimated Impact on Violence	1	10 (United States)
Contributing to EDIE Information	0	6 (United States)
Contributing to Implementation	5	9 (United States)
Contributing to Cost Data	2	1 (United States)

Appendix 3. Characteristics of included studies for effectiveness

Authors (Year)	Country	Study Design	Intervention	Population/ Place	Comparison	Outcomes Measured	Quality Level	Findings
Baker & Wolfer (2003)	US	Quasi-Experimental Design	POP approach involving target hardening, target-oriented patrols / proactive patrols, school visits, curfew, crackdown on drug and alcohol use in public, and establishment of a Neighbourhood Watch.	Community members living in the vicinity of a local park (wave 1 N = 124; wave 2 N = 125).	People who live in the general borough population, with access to the park, but do not live in the immediate vicinity (wave 1 N = 337; wave 2 N = 333).	Victim of crime Community Connectedness	High	Fall in victimization from the pre- to post-assessment is greater for the target group (d = −0.33) than the control group (d = −0.17). Incidents of drug or disorderly behaviour more than halved for the target group. Both the target and control groups felt significantly safer in the park both during the

								day and at night by the post-assessment (p<.001). The target group felt significantly safer than the control group during the day at post-assessment (p<.05), whilst there was no difference between the target and control at night.
Beito & Sigler (1997)	US	Quasi-Experimental Design	Community Service Officers Programme: Two officers permanently assigned to the area established a working relationship with residents through community	A high crime neighbour-hood in Tuscaloosa, Alabama, including a low-income housing project with one of the highest	Comparison neighbour-hood with a low-income housing project, similar population density, geographic and	Crime count data	Moderate	Overall increase of 16.6% in crime count data for the experimental condition and an overall increase of +2.0% for the control condition; 6.4% of residents in the experimental neighbourhood reported being a victim

			meetings and a community office, promoted self-help techniques e.g. neighbourhood watch, property identification programmes, referred problems to social services agencies when appropriate, and promoted a general community improvement effort	crime rates in the city	demographic characteristics			of crime in the past 18 months Vs 6.7% of the control group's residents
Boston Police Departm	US	Other quantitative design	POP approach: Analysis of the problem to map incidents and	Residents in the Brighton/ Allston area of Boston,	No comparison group – pre/post	Breaking and	Moderate	Overall decrease in the number of breaking and entering incidents, from 650 in 2006 to 390

ent (2008)			identify trends, risk factors, and repeat offenders; Response involved the increased use of police resources with a new protocol for breaking and entering calls, creation of a 'Watchlist' of known repeat offenders, community education programme with local building owners to encourage target hardening to	the district with the most residential burglaries in the city in 2006	comparison within the district only	entering rates		in 2007 (decrease of 40%). Decrease in the rate of multiple breaking and entering incidents at a single address, from 66% of incidents in 2006 to 16% of incidents in 2007.

			reduce risk factors, targeting of high-risk 'hotspots' with increased police patrols and operations targeting high-risk times such as Spring Break					
Braga et al. (2001)	US	Quasi-experimental design	Operation Ceasefire: A direct law-enforcement attack on illicit firearms traffickers supplying young people with guns and an attempt to generate a strong deterrent to gang violence through	Young people who are members of gangs in Boston	Other US and New England cities	Youth homicides, gun assaults, shots fired, and youth gun assaults	Moderate	63% decrease in the monthly number of youth homicides; 25% decrease in the monthly number of citywide gun assault incidents: 32% decrease in the monthly number of citywide shots-fired calls for service; 44% decrease in the

			a 'pulling levers' strategy					monthly number of youth gun assaults
Braga (2008b)	US	Quasi-experimental design	Operation Peacekeeper: An interagency task force targeting gang-involved young people to prevent gun violence through a 'pulling levers' strategy – responding to any gang gun violence with a law enforcement 'crackdown', ongoing communication through gang forums, offering services and	Young people involved in gangs in Stockton, California	Other similar-sized cities in California	Gun homicide rates	Moderate	42% decrease in the monthly number of gun homicide incidents ($p = 0.045$) None of the comparison cities experienced a statistically significant reduction in the monthly count of gun homicides that coincided with the implementation of the Peacekeeper intervention in Stockton

			opportunities to gang members as an alternative to violence					
Braga et al. (2008)	US	Quasi-experimental design	Project Safe Neighbourhoods Initiative: Interagency taskforce implementing 'pulling levers' strategy in response to gang violence - focusing prevention, intervention, and enforcement activities on gang members	Young people who are members of gangs in Lowell, Massachusetts	Other major Massachusetts cities	Gun homicides and gun-aggravated assault incidents	Moderate	27.8% decrease in the mean monthly count of gun homicides and gun assault incidents 43.1 percent decrease in the monthly number of assaultive gun violence events (p = 0.002)

			involved in violent conflicts					
Braga & Schnell (2013)	US	Quasi-experimental design	Boston Safe Streets Team (SST): 13 violent crime hot spots identified and assigned officers. In each hotspot, the team identified recurring problems and implemented 396 different strategies grouped into environmental interventions, enforcement interventions, and community	13 areas identified as violent crime hot spots in Boston, Massachusetts	Street segments at least two blocks away from intervention streets that were not receiving the SST intervention	Yearly violent index crimes in a street unit	High	Statistically significant 17.3% reduction ($p < 0.05$) in the number of violent index crime incidents; statistically significant 19.2% reduction ($p < 0.05$) in the number of robbery incidents, and a statistically significant 15.4% reduction ($p < 0.05$) in aggravated assault incidents at the treatment street units relative to the comparison street units.

			outreach/social service interventions					
Gulf Breeze Police Department (2004)	US	Pre/post comparison	Seminars for parents and teachers to educate them about drugs and drug activity among young people in the community	Adolescents (ages 10-17) attending Gulf Breeze High School	N/A – pre/post comparison only	Alcohol and drug offences	Low	42.3% reduction in misdemeanour drug possession offences from 2002 to 2003 66.7% reduction in alcohol offences from 2002 to 2003
Lancashire Constabulary (2008)	UK	Quasi-experimental design	Enforcement responses such as ASBOs, drug warrants; target hardening and estate clean up; social crime prevention such as restorative justice for criminal	Residents of the Farringdon Park neighbourhood	Average crime count in the two years prior to the intervention	Crime; calls for service; burglary; criminal damage	High	From 2004/5 to 2007: 57% reduction in crime 46% reduction in calls for service 55% reduction in burglary

			damage offenders, positive community activities, referral to youth services					57% reduction in criminal damage 37% reduction in anti-social behaviour
Maguire et al. (2003)	US	Pre/post comparison	A School Resource Officer worked with students and staff at a middle school to determine what issues students were most concerned about and to help them develop solutions such as peer mediation and increased teacher presence in the corridors	Students at Thurston Middle School in Westwood, Massachusetts	N/A – pre/post comparison	Experience of bullying or intimidation Level of trust in the School Resource Officer	Moderate	No significant change in experience of bullying or intimidation Significant decrease in trust in the School Resource Officer

Mazerolle et al. (1998)	US	Quasi-experimental design	Each housing site had an interagency 'site team' which met every month to discuss crime problems and coordinate problem-solving strategies; strategies included referral to treatment and counselling programmes for residents, improving lighting, and traditional law enforcement e.g. arrests, surveillance etc.	Six public housing sites in Jersey City experiencing severe and persistent crime problems	N/A – no comparison group	Citizen calls for police service	High	25% decrease in calls for service across the six public housing sites receiving the intervention compared to a 4% decrease city wide over a 2.5 year period

Preston Early Intervention Partnership (2012)	UK	Quasi-experimental design	Home visits, a tour of the custody complex, and referral to diversionary activities and other services as needed	Young people involved in anti-social behaviour identified by police while on patrol and entered into the Youth Referral Database	Rates of arrest for the first 60 young people entered into the Youth Referral Database in each of the two years preceding the intervention (120 in total)	Arrest	Moderate	Out of the 120 young people participating in the intervention, 24 were arrested compared to 65 out of the 120 young people in the comparison group, indicating 41 arrests were avoided
Stokes et al. (1996)	US	Quasi-experimental design	Scanning - identified four middle schools and solicited participation from the principals; gathered data and information on the context of	Middle school students in North Philadelphia neighbourhoods – intervention took place in	Comparison group was three other schools in the same area, receiving 'treatment as usual'	Community connected ness – fear of being attacked; Victim of crime – being bothered or	Moderate	No significant change in victimisation ($F = 0.0707$, $p = 0.7904$) or fear of being attacked ($F = 0.0740$, $p = 0.7857$) in the intervention school; significant decrease in the proportion of students

			the problem of student victimization to and from school Analysis - used focus groups, a victimization survey, and police and school data to get a better picture of student victimization and to create maps of incidents / unsafe areas around each school Response - chose the test school, selected the 'Safe Corridor', patrolled the safe corridor twice a day in	John Wanamaker Middle School, with 846 students, 90% Black and 10% Hispanic		attacked during journey to or from school		being attacked (F = 20.0775, p<0.0001) and no significant change in the fear of being attacked in the control group (F = 0.3702, p = 0.5429) In the intervention school, only 27.4% of students knew about the corridor and only 25.3% used the corridor.

			morning and afternoon					
Taylor et al. (2024)	US	Randomised controlled trial	POP for Youth: Training in POP and community policing strategies, as well as youth development and strategies to improve police-youth interactions; patrol squads visited assigned hot spots at least once per day and implemented at least one POP for Youth project in each hot spot	Adolescents (ages 10 to 17) living in 128 crime hot spots identified across three mid-Atlantic cities	Hotspots were randomly allocated to receive either the intervention or standard patrol	Property crime and violent crime counts	High	No significant effect on violent crimes Some evidence of an unintended increase in property crime in intervention hot spots in two of the cities

Transport for London (2008)	UK	Pre/post comparison	Problem oriented approach to reducing violence involving young people and antisocial behaviour on bus network: behaviour code for young people, implementing vandal-resistant materials in bus shelters/windows, working with schools to provide alternative travel routes, teachers at bus stops, police patrols at hot spots, bus CCTV, crime	Adolescents 10-17 years travelling on the London bus network	N/A – pre/post comparison only	Rate of crime involving young people on bus passenger journeys, including violence against the person and robbery Driver reports of antisocial behaviour Criminal damage incidents	Low	The rate of crime involving young people per million bus passenger journeys was 24 in 2005/6 and 23 in 2006/7 (a 4% improvement). The rate of crime involving young people fell to 16 in 2007/8 (a 33% improvement on 2007/8). Crime on the bus network allegedly involving one or more under 16 suspects in 2007/8 was 24% lower (1,431 offences) than in 2006/7 when it peaked. Robbery fell by 31% (465 incidents),

			prevention posters, focused work with young people who offended on buses, presentations at schools, driver training					violence against the person fell by 3% (40 incidents) and criminal damage fell by 31% (649 incidents). Number of driver reports of antisocial behaviour fell by 6% (2,938 incidents) between 1/09/2007 and 30/04/08) compared to the same period the previous year. Criminal damage incidents fell by 27% (1,583 incidents).
Weisburd et al. (2008)	US	Randomised controlled trial	Redlands Risk-Focused Policing at Places (RFPP) Program: Interventions	High school students living in census blocks in	Census blocks assigned to the comparison	26 outcome variables across community,	Moderate	No statistically significant differences between intervention & control group apart from a marginally

			included community activities, parenting classes, neighbourhood meetings, neighbourhood clean-ups, recreation activities, tutoring	Redlands, California, identified as having high risk factors for crime	condition received standard police patrols	family, school, peer-individual domains, plus self-reported delinquency, drug use, and arrest, and procedural justice		significant effect on the average level of arrest for the intervention group.

Appendix 4. Measured outcomes across included studies for effectiveness

YEF outcome framework category *Category description (as described in the framework)*	Measured outcomes *(descriptions by study authors)*	Studies
Breaking the Law or Offending Behaviour *Focuses on criminal behaviours ranging from minor offences that do not involve violence against another person to serious violent crimes which do.*	Arrest (self-reported), Arrest (police statistics), Delinquency scale (self-reported), Violent index, Homicide, Rape/sex assault, Robbery, Aggravated assault, Rate of youth crime per million bus passenger journeys, Incidents of etching on buses, Incidents of graffiti on buses, Total count of all Part I property crimes per month, Calls for police service, Misdemeanour drug possession offences, Minors in possession of alcohol offences, Youth homicides (controlling for trends and seasonal effects), Gun assaults (controlling for trends and seasonal effects), Shots fired (controlling for trends and seasonal effects), B-2 Youth gun assaults (controlling for trends and seasonal effects), Youth homicides (controlling for additional factors), Gun assaults (controlling for additional factors), Shots fired (controlling for additional factors), B-2 Youth gun assaults (controlling for additional factors), Gun homicides, Violent crime (gun homicides and gun-aggravated assaults), Residential breaking and entering crimes by year, Simple assault,	(N=13) Beito & Sigler (1997); Boston Police Department (2008); Braga (2008b); Braga et al. (2001); Braga et al. (2008); Braga et al. (2012); Gulf Breeze Police Department (2004); Lancashire Constabulary (2008); Mazerolle et al. (1998); Preston Early Intervention Partnership (2012); Taylor et al. (2024); Transport for London

	Burglary, Theft, Vehicle theft, Total crime, Anti-social behaviour, Criminal damage	(2008); Weisburd et al. (2008)
Community Connectedness *Feeling connected to one's community can look like identifying as a member of the community, trusting others in the community, having personal needs met by the community and engaging in community activities.*	Community disorganisation, Low attachment, Rewards for pro-social involvement, Perception of procedural justice, Fear of being attacked, Feel safe in the park during the day, Feel safe in the park during the night, Feel safe due to crime prevention efforts	(N=3) Baker & Wolfer (2003); Stokes et al. (1996); Weisburd et al. (2008)
Victim of Crime *Also called 'criminal victimisation'. Having experience, or being a victim, of different types of crime, including robbery, theft, vandalism, assault and kidnapping.*	Were you ever bothered or attacked on your way to or from school? Had property broken into in past 6 months, Had something stolen from home in past 6 months? Noticed vandalism in past 6 months? Noticed public drinking/disorderly conduct in past 6 months? Was ever the victim of a crime? Have you ever been bullied or teased to the point you felt unsafe or unwelcome at school? Have you ever been hit, kicked, punched, tripped, pinched or something like that during the school year?	(N=3) Baker & Wolfer (2003); Stokes et al. (1996); Maguire et al. (2003)
Family Relationships and Support *Positive and supportive relationships with family members.*	Family attachment, Poor family supervision, Family attitudes favourable to antisocial behaviour, Family opportunities for pro-social involvement, Family rewards for pro-social involvement	(N=1) Weisburd et al. (2008)

Drug and Alcohol Use *Also called 'substance misuse/abuse'. Problematic use of drugs and/or alcohol that results in negative and harmful consequences to the self or others, such as impaired physical health, difficulties concentrating or skipping school.*	Attitudes favourable to drug use, Perceived risk of drug use, Drug use scale (self-reported)	(N=1) Weisburd et al. (2008)
School Engagement *School engagement is a multifaceted construct including affective, behavioural, and cognitive components. The sub-outcomes below are different indicators of school engagement.*	School failure, Low commitment to school	(N=1) Weisburd et al. (2008)
School Environment *Also called 'school climate'. A positive whole-school culture encompasses emotional, relational and physical*	Opportunities for pro-social involvement, Rewards for pro-social involvement	(N=1) Weisburd et al. (2008)

safety, as well as cultivates shared nurturing values. This environment fosters the wellbeing of children, young people, parents/carers and staff.		
Positive and Prosocial Identity *Viewing yourself as someone who engages in positive and meaningful activities and not in criminal activities*	Rebelliousness, Attitudes favourable to delinquency/antisocial behaviour	(N=1) Weisburd et al. (2008)
Criminal peers *Also called 'delinquent* peers' Having a close group of people who take part in and promote criminal behaviour – criminal behaviour may be an important part of the group's identity.* **The term 'delinquent' is falling out of use to reduce stigma*	Interaction with antisocial peers, Rewards for antisocial behaviour	(N=1) Weisburd et al. (2008)

Appendix 5. Characteristics of included studies for implementation

Authors (Year)	Country	Study Design	Intervention	Quality Level	Implementation Outcomes	Experiences of CYP
Andrews (2024)	UK	Process Evaluation	Deter and detect approach, involving problem-oriented and intelligence-led tactics to reduce knife crime (e.g., police cadets testing purchases of bladed articles, stop-and-searches of known habitual knife carriers)	Very Low	**Acceptability:** Approach was perceived as acceptable to the community, with areas in which police were present determined by concerns of the community. Visibly tackling a community concern increases legitimacy. Any stop-and-search strategies were suggested as acceptable as they were intelligence led rather than targeted at any demographic. **Appropriateness:** Deployment of the intervention team was based on superficial analysis of command areas, with the use of hot spots recommended for future implementation.	N/A

					The rate of 'positive outcomes' (i.e., searches resulting in finds) was consistently higher than the national average, highlighting that any stop-and-search strategies were appropriately used.	
Beito & Sigler (1997)	US	Quasi-Experimental Design	Community Service Officers Programme: Two officers permanently assigned to the area established a working relationship with residents through community meetings and a community office, promoted self-help techniques (e.g. neighbourhood watch), property identification programmes, referred problems to social	Moderate	**Acceptability:** Community members in the experimental condition had more positive attitudes toward police efforts to reduce crime, and a reduction in pathology, than those in the control condition, suggesting they benefitted from the presence of community service officers. **Adoption:** Social services agency staff highlighted that should the programme be expanded, there would need to be additional resources given an expected increase in the flow of clients. **Appropriateness:** There was a lack of theoretical basis concerning the development of the intervention. In addition, there was a lack of information surrounding how officers made	N/A

			services agencies when appropriate, and promoted a general community improvement effort		decisions within the intervention, regarding what should/should not be targeted and how.	
Braga et al. (2008)	US	Quasi-experimental design	Project Safe Neighbourhoods Initiative: Interagency taskforce implementing 'pulling levers' strategy in response to gang violence - focusing prevention, intervention, and enforcement activities on gang members involved in violent conflicts	Moderate	**Adoption:** POP strategies were tailored to particular gangs. For Asian gangs, gang 'elders' and their illegal gambling businesses were targeted directly as these were more likely to be known to the police. For other gangs, young people themselves were targeted. The degree of intervention varied according to how embedded individuals were within the gang. For many, heightened levels of local police patrols and probation were sufficient to end violence, whilst other hardcore members needed enhanced enforcement from federal authorities. **Appropriateness:** The interagency taskforce was confident that POP was appropriate to targeting Hispanic gangs. They were less	N/A

					confident in their ability to effectively target Asian gang violence, as they are more organised and secretive in nature.	
Braga & Schnell (2013)	US	Quasi-experimental design	Boston Safe Streets Team (SST): 13 violent crime hot spots identified and assigned officers. In each hotspot, the team identified recurring problems and implemented 396 different strategies grouped into environmental interventions, enforcement interventions, and community outreach/social service interventions	Moderate	**Appropriateness:** Most staff were adept at using crime mapping software, with additional support from analysts available, enabling appropriate levels of problem-oriented responses according to levels of crime in the area. However, some relied on less nuanced analysis, including informal assessments based on their experiences in the area, which may have impacted on appropriate use of problem-oriented responses in these areas. **Fidelity:** 13 teams implemented an average of 30.5 POP interventions per identified hotspot. The number of interventions implemented varied per area, dependent upon leaders' commitment to a POP approach. For low-fidelity areas, command staff had to be provided with additional training and closer supervision.	N/A

Bullock et al. (2023)	UK	Process Evaluation	Explored 44 police personnel's experiences of implementing problem-oriented policing to target knife crime.	Moderate	**Acceptability:** Commitment to a problem-oriented approach has not permeated throughout the whole police force. Engaging police in prevention strategies remains challenging. **Adoption:** Knowledge of the violence hotspots (including transport hubs) and times of day that criminality takes place is key to implementing an intervention effectively. Flexibility is required, with interventions needing to be adapted according to the context and location. **Appropriateness:** Police personnel flagged that the context is critical, highlighting that POP may be appropriate in some areas (such as London) but less so in other, out-of-city locations. **Feasibility:** Difficulties analysing knife crime data due to availability of analysts, quality of data, reliance on limited and potentially biased police data, and difficulties accessing data	N/A

					from other organisations (e.g., emergency health services) impact on the effective use of POP. POP provides a *"blanket wrap[around] support"* for communities, but it would take a lot of work to support communities to problem solve and empower them to do so, which may affect feasibility. **Fidelity:** Police personnel reported that there is a lack of evaluation conducted on any operation put into place, so it is challenging to understand whether the approaches have been fully adhered to. **Reach/Penetration:** Most police forces across England and Wales have, at some point, used POP. However, this is inconsistently and unevenly applied across the country. **Sustainability:** Cost of sustaining delivery over time was highlighted as particularly problematic.	

Glover (2002)	US	Process Evaluation	Examines community and POP strategies used by US police departments to target school violence, highlighting process dimensions crucial for successful implementation.	Very low	**Acceptability:** Utilising policing strategies in schools can be problematic with school and police departments perceived as incompatible organisations by both members of each organisation and the wider public, impacting on their perceived acceptability. This is particularly problematic when a distinction needs to be made between a crime and a disciplinary problem. **Adoption:** For a POP approach to be effectively adopted, this requires a collaborative approach to its design. Targeting school violence requires close partnerships between schools and police, effective implementation planning, clear and timely feedback between organisations, managerial support for the programme, information management, and use of evidence-based approaches. **Appropriateness:** Schools have several characteristics that make them appropriate areas to target for interventions. This includes	N/A

					having a relatively stable membership, defined geographical location, clear operating hours, clear social structures and roles, defined resources, high visibility in the community and acceptance as a legitimate institution. **Feasibility:** When considering implementing POP in a school, it is important to differentiate between community- and school-based interventions for this to be feasible. School administrators considering implementing community-based approaches need clarity on the extent to which any approach extends beyond the boundaries of the school and the school environment.	
Gulf Breeze Police Depart-ment (2004)	US	Pre-post Design	Seminars for parents and teachers to educate them about drugs and drug activity among young people in the community	Low	**Acceptability:** Many parents did not perceive there to be a problem with drugs within the city or with their child, which caused some difficulty in getting parents to attend seminars. Teachers provided positive feedback to officers for their initiative and information provided in the seminars.	N/A

					Adoption: Getting parents to attend seminars caused some difficulty in adopting this approach. POP was adopted well by all involved, including school resource officers, canine officers, command staff, and police chief. Adoption was supported by school administrators, who encouraged teachers to attend training and provided classroom and media support. Teacher participation was very high and well received. Media publicised the drug and alcohol problem in the community, encouraging parents and members of the public to attend training. Support from local faith organisations led to the provision of facilities for meetings, as well as publicising the training through congregations. **Appropriateness:** To demonstrate that the programme was appropriate and applicable to local residents, training resources on types	

					of drugs, effects, and distribution, were bought back to focus on the local community; highlighting how those drugs are a problem in the particular area. A collaborative approach to developing a comprehensive plan to tackle drug and alcohol use in the local community was taken, with law enforcement, school administrators, active citizens, media, teachers and student representatives helping with its creation. This led to a plan which was comprehensive, informative and appropriate for the local area. **Feasibility:** The programme was perceived as feasible to implement, with only 64 overtime hours needed across the course of the programme. The school and police planned to continue the programme with new classes each year. **Sustainability:** Authors reported that it would take several years of implementing this approach to control drug and alcohol use, but	

					that it remained a sustainable approach. It would have to be embedded within ongoing annual operations, with each new school group targeted separately. **Cost:** The programme providers were successful in securing additional funding from a Local Law Enforcement Block Grant to fund the overtime for bicycle patrols. The main costs were personnel and time, but given the positive results, the additional costs and resources were perceived as reasonable and justified.	
Lanca-shire Constab-ulary (2008)	UK	Quasi-experimental design	Implemented evidence-based enforcement, situational and social crime prevention strategies, using a SARA approach.	High	**Adoption:** Some time-slippage was experienced in the planned implementation and adoption of POP, particularly regarding implementing situational crime prevention responses (i.e., making physical changes in the area) and court processing of enforcements. **Cost:** There were considerable cost savings, including $62,100 associated with a reduction in burglary, $72,420 due to a reduction in	N/A

					criminal damage, and $51,770 related to a reduction in antisocial behaviour.	
Maguire et al. (2003)	US	Pre-post Design and process evaluation	Implemented a schools-based partnership with local police in Westwood. In collaboration with students, strategies devised to reduce any problematic behaviour included clear school rules that were uniformly and consistently enforced, increased presence of teachers in hallways, and utilising peer mediation to overcome disputes	Moderate	**Acceptability:** The use of multi-agency planning and input from young people increased the acceptability of the programme. Exposure to police officers through this collaborative approach led to the development of positive relationships between police, teachers, and young people. **Adoption:** Changes in staffing affected the continued adoption of this approach temporarily. New staff also implemented additional programmes distinct from POP, which may have impacted upon the findings. Despite this, young people engaged well in problem-solving training sessions provided as part of the POP approach.	N/A

Mazerolle et al. (1998)	US	Quasi-experimental design	Each housing site had an interagency 'site team' which met every month to discuss crime problems and coordinate problem-solving strategies; strategies included referral to treatment and counselling programmes for residents, improving lighting, and traditional law enforcement e.g. arrests, surveillance etc.	High	**Adoption:** Completing a high number of problem-solving activities, alongside continued engagement in problem-solving teams, contributed to positive outcomes, including a reduction in calls for help related to serious crime.	N/A
Preston Early Intervention	UK	Quasi-experimental design	Implemented POP in Lancashire. Strategies used included early intervention analysis,	Moderate	**Cost:** There was a cost saving due to a reduction in arrests, equating to £82,000 (£130,000 in 2008/9 vs. £48,000 in 2010/11).	45% of CYP said they felt 'scared' or 'terrible' following the custody visit,

Partnership (2012)			home visits to identify risk factors and offending level assessments, utilising a custody experience visit to educate and prevent arrests of young people, and creating bespoke multi-agency action plans to prevent offending.			with a further 31% feeling 'cold and lonely', 'worried' or 'sad'.
Taylor et al. (2024)	US	RCT	Examined whether POP in hot spots of crime reduces poverty and violent crime in three cities. This study compared POP vs. control hot spots.	High	**Adoption:** POP was not well-adopted by officers, with few visits to hotspot areas conducted per month. On visits to hotspot areas, challenges in engaging community members were observed. Authors suggested that the after-effects of COVID-19 and strained police relationships related to the George Floyd murder stifled police interactions with young people.	N/A

					Fidelity: Low fidelity was observed, with between 1.3-7.4 visits to hotspots per month, compared to the recommended 16-17 visits.	
Transport for London (2008)	UK	Pre-Post Design	Compared the effectiveness of a problem-oriented approach to reducing violence involving young people (primarily targeting adolescents) on the bus network in London, pre- and post-implementation. Authors provided some feedback on implementation challenges.	Low	**Acceptability:** Public perceptions and media discourse on increased crime levels enhanced the acceptability of focused efforts to tackle the problem. **Adoption:** Challenges included the political and media pressure for a quick resolution to violence involving young people on the bus network. It was difficult to gain useful intelligence at a local level to support implementation in a timely manner. Transport for London received inconsistent support or push back from schools and scepticism and resistance from some transport staff responsible for implementing parts of the initiative.	N/A

					Feasibility: The number of young people making journeys on the bus network was high (accounting for 1,100,000 journeys a day on the bus network), with free travel for under 16s contributing to a 42% increase in antisocial behaviour. With such a large volume of young people, it would be challenging to rollout interventions that can target all potential users. The rollout of some of the interventions was planned to be gradual, which gave time to overcome and modify the intervention as issues were identified. **Fidelity:** The fidelity of POP across the transport network was impacted by resistance from some employees and inconsistent support across schools, meaning the approach was inconsistently applied.	
Weisburd et al. (2008)	US	RCT	Evaluated the effectiveness of the Risk-Focused Policing at Places approach; a	Moderate	**Appropriateness:** Police took a 'one size fits all approach' for implementing interventions in the community, inconsistent with POP. The use	N/A

			community-based and problem-oriented approach targeting risk and protective factors related to delinquency and problem behaviours of young people living in census block group areas. Authors provided feedback on the implementation of the intervention.		of large geographic areas may have prevented targeted interventions. **Fidelity:** The 'dosage' (degree of contact between community and police) may not have been sufficiently high to lead to positive outcomes.	

Appendix 6. Availability of evidence according to each of Proctor et al.'s (2011) implementation outcomes

Authors (Year)	Acceptability	Adoption	Appropriate-ness	Feasibility	Fidelity	Reach/ penetration	Sustainability	Cost
Andrews (2024)	Yes	No	Yes	No	No	No	No	No
Beito & Sigler (1997)	Yes	Yes	Yes	No	No	No	No	No
Braga et al. (2006, 2008)	No	Yes	Yes	No	No	No	No	No
Braga & Schnell (2013)	No	No	Yes	No	Yes	No	No	No
Bullock et al. (2023)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Glover (2002)	Yes	Yes	Yes	Yes	No	No	No	No
Gulf Breeze Police Department (2004)	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Lancashire Constabulary (2008)	No	Yes	No	No	No	No	No	No
Maguire et al. (2003)	Yes	Yes	No	No	No	No	No	No

Mazerolle et al. (1998)	No	Yes	No	No	No	No	No	No
Preston Early Intervention Partnership (2012)	No	Yes	No	No	No	No	No	Yes
Taylor et al. (2024)	No	Yes	No	No	Yes	No	No	No
Transport for London (2008)	Yes	Yes	No	Yes	Yes	No	No	No
Weisburd et al. (2008)	No	No	Yes	No	Yes	No	No	No