



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/191072/>

Version: Accepted Version

---

**Article:**

Case, S, Sutton, CE, Greenhalgh, J et al. (2022) Searching for context: a review of "what works" reviews of interventions to prevent youth offending using the EMMIE Framework. *Safer Communities*, 21 (4). pp. 272-289. ISSN: 1757-8043

<https://doi.org/10.1108/SC-04-2022-0014>

---

© 2022, Emerald Publishing Limited. This is an author produced version of an article published in *Safer Communities*. Uploaded in accordance with the publisher's self-archiving policy.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



**Contextualising evaluations of interventions to prevent youth offending: 'What Works' and EMMIE**

Journal:	<i>Safer Communities</i>
Manuscript ID	SC-04-2022-0014.R2
Manuscript Type:	Research Paper
Keywords:	EMMIE, effectiveness, children, prevention, youth justice, What Works, Realist Synthesis, risk

## MANUSCRIPT DETAILS

TITLE: Contextualising evaluations of interventions to prevent youth offending: â€˜What Worksâ€™ and EMMIE

## ABSTRACT:

To examine the extent to which â€˜What Worksâ€™ reviews in youth justice enable understanding of the features of effectiveness (what works, for whom, in what circumstances and why?) specified in the EMMIE framework.

The EMMIE framework examined findings within a sample of â€˜What Worksâ€™ style reviews of preventative youth justice intervention effectiveness.

â€˜What Worksâ€™ style reviews of evaluations of preventative youth justice interventions often omit the requisite details required to examine all of the necessary elements of effectiveness contained within the EMMIE framework. Whilst effectiveness measures were typically provided, the dominant evaluation evidence-base struggles to consider moderators of effect, mechanisms of change, implementation differences and cost effectiveness, Therefore, â€˜What Worksâ€™ samples cannot facilitate sufficient understanding of â€˜what works for whom, in what circumstances and why?â€™.

CUST\_RESEARCH\_LIMITATIONS/IMPLICATIONS\_\_(LIMIT\_100\_WORDS) :No data available.

CUST\_PRACTICAL\_IMPLICATIONS\_\_(LIMIT\_100\_WORDS) :No data available.

CUST\_SOCIAL\_IMPLICATIONS\_\_(LIMIT\_100\_WORDS) :No data available.

We extended the approach adopted by an earlier review of effectiveness reviews (Tompson et al, 2020), considering more recent reviews of the effectiveness of preventative interventions using the EMMIE framework. Unlike previous reviews, we prioritised the utility of the EMMIE framework for assessing the factors affecting the effectiveness of preventative interventions in youth justice.

## Searching for context A review of ‘What works’ reviews of interventions to prevent youth offending using the EMMIE framework

### Structured abstract

- **Purpose:** To examine the extent to which ‘What Works’ reviews in youth justice enable understanding of the features of effectiveness (what works, for whom, in what circumstances and why?) specified in the EMMIE framework.
- **Design/methodology/approach:** The EMMIE framework examined findings within a sample of ‘What Works’ style reviews of preventative youth justice intervention effectiveness.
- **Findings:** ‘What Works’ style reviews of evaluations of preventative youth justice interventions often omit the requisite details required to examine all of the necessary elements of effectiveness contained within the EMMIE framework. Whilst effectiveness measures were typically provided, the dominant evaluation evidence-base struggles to consider moderators of effect, mechanisms of change, implementation differences and cost effectiveness. Therefore, ‘What Works’ samples cannot facilitate sufficient understanding of ‘what works for whom, in what circumstances and why?’. We argue that Realist Synthesis can fill this gap and shed light on the contexts that shape the mechanisms through which youth justice interventions work.
- **Originality:** We extended the approach adopted by an earlier review of effectiveness reviews (Tompson et al, 2020), considering more recent reviews of the effectiveness of preventative interventions using the EMMIE framework. Unlike previous reviews, we prioritised the utility of the EMMIE framework for assessing the factors affecting the effectiveness of preventative interventions in youth justice.

**Keywords:** youth justice; Realist Synthesis, what works; intervention; evaluation; EMMIE.

## Introduction

In this article, we assert that artificially inflating and misrepresenting the ‘effectiveness’ of preventative interventions in the youth justice field through over-reliance on the restricted ‘What Works’ evaluation framework has provided the sector with a perfect storm of reductionism and decontextualisation when understanding and responding to the lives of children who offend. We extrapolate this critique by ultimately advocating for a contextualised Realist Synthesis approach to the evaluation of youth crime prevention initiatives, supporting the development of a more valid and complete evidence-base of ‘effectiveness’ to better support youth justice practice.

The traditional definitions of intervention ‘effectiveness’ in youth justice derive from legislation, notably the Crime and Disorder Act (1998), wherein ‘effectiveness’ measures are inextricably linked to the prevention and reduction of offending through a focus on individualised/psychosocial, responsabilising (holding the child primarily responsible for their actions) and adulterising (adult-centric and defined) ‘risk factors’. Since 1998, the ‘What Works’ framework (cf. Sherman et al 1998; Cochrane Collaboration 2011) has dominated the assessment and evaluation of preventative interventions in the Youth Justice System of England and Wales – a risk-based framework that utilises (quasi) experimental, quantitative methodologies to measure and determine intervention ‘effectiveness’. However, the reductionist, partial (biased and incomplete) and psychosocial/individualised, risk-led assessments of and responses to offending by children perpetuated and validated by the ‘What Works’ approach have fostered the *decontextualisation* of interventions in youth justice (Sutton et al 2022) – a limitation acknowledged by several of its key proponents<sup>1</sup>. For example, a recent review of the evidence-base for crime reduction programmes identified that ‘Initially, the UK What Works Centers largely followed [‘What Works’] on what constituted good evidence’ (Tompson et al 2002: 2). However, ‘the more systematic analysis of what is working where, and why – is rapidly becoming the new normal. (Halpern, 2018: 4), largely because ‘it is considered helpful to make clear how outcome results might vary according to the context in which an intervention is to be implemented’ (Tompson et al 2020: 2).

Therefore, the reductionist, decontextualising limitations of the established ‘What Works’ evaluation framework in youth justice and the responsabilising and adulterising excesses of the risk-based interventions that it privileges (see Case 2021) strongly indicate that research and practice should explore alternative methodologies for conceptualising and evaluating the ‘effectiveness’ of preventative interventions. Additionally, the rapidly emerging context of ‘Child First’ justice based on the tenets<sup>2</sup> of seeing children as children, promoting prosocial identity for positive outcomes,

---

<sup>1</sup> For example, in their highly influential meta-evaluation of crime prevention programs, ‘Preventing Crime: What Works, What Doesn't, What's Promising’ (Sherman et al (1998) acknowledged that ‘What’s Unknown’ in terms of intervention effectiveness is heavily influenced by variations in ‘social setting’ (i.e. context). Further, a UK government-commissioned systematic Rapid Evidence Assessment of more qualitative, context-sensitive youth crime interventions underpinned by ‘What Works’ concluded that the framework is a potential to exclude effective interventions due to its over-reliance on evaluation quality rather than intervention quality (Adler et al 2016).

<sup>2</sup> *See children as children*: Prioritise the best interests of children, recognising their particular needs, capacities, rights and potential. All work is child-focused, developmentally informed, acknowledges structural barriers and meets responsibilities towards children;

*Develop pro-social identity for positive child outcomes* Promote children’s individual strengths and capacities to develop their pro-social identity for sustainable desistance, leading to safer communities and fewer victims. All work is constructive and future-focused, built on supportive relationships that empower children to fulfil their potential and make positive contributions to society;

*Collaborate with children* Encourage children’s active participation, engagement and wider social inclusion. All work is a meaningful collaboration with children and their carers;

collaboration with children and diversion from the formal Youth Justice System (YJB 2021), is supplanting established (reductionist, decontextualised) understandings of and responses to children who offend. Formalising Child First represents a significant policy and strategy shift for youth justice in England and Wales - a deliberate move away from the hegemonic risk management model, which has been criticised as inappropriate, criminalising and iatrogenic due to its privileging of the prevention of negative behaviours/outcomes (rather than promoting positive behaviours/outcomes), neglect of (Child First) meaningful participation, prioritisation of adult-centric understanding/practices and tendency towards ‘interventionism’ (see Case and Haines 2021). Critiques of risk management practice models and mechanisms (e.g. risk assessment) are comprehensive, longstanding and evidence-based - focused on theoretical, conceptual, methodological, practical and ethical weaknesses/limitations<sup>3</sup> (cf. Case and Hazel, in press; Case and Haines 2009). Therefore, the implications of this conceptual and practical move from risk management to Child First youth justice is immense and further recommends an urgent reconceptualisation and contextualisation of the content, delivery and evaluation of youth justice interventions traditionally shaped by risk-based understandings of offending and evaluated using the experimental ‘What Works’ framework. Indeed, the evolution of Child First as a policy and strategic imperative (YJB 2021) into a consistently understood and implemented practice reality is arguably the most pressing issue facing youth justice in England and Wales (cf. Case and Hazel, in press; Case and Browning 2021a,b; Day 2022) - one currently shaping forthcoming revisions to Case Management Guidance to practitioners and YJB discussions with the inspectorate regarding more holistic evaluation of Youth Offending Team practice<sup>4</sup>.

A promising alternative, contextualised and Child First-compatible mode of intervention evaluation is offered by *realist* evaluation<sup>5</sup> (Pawson, 2013, 2006a, Emmel et al, 2018), which directly and explicitly challenges the perceived limitations of ‘What Works by adopting a more qualitative, multifaceted approach to conceptualising and understanding intervention effectiveness that considers intervention *outcome* measures alongside essential features of the intervention *context* (e.g. implementation differences) and the *mechanisms of change* activated during an intervention (cf. Sutton et al 2022). For example, context can be understood in realist terms as the relational and dynamic features that shape the mechanisms by which an intervention works<sup>6</sup> (Greenhalgh and Mazano 2021) – implying a potential alignment with Child First tenets of collaborative and participatory (relationship-based rather than risk-led) practice seeking to promote prosocial identity (as a key mechanism of change) and positive outcomes (Case and Browning 2021a).

---

*Promote diversion:* Promote a childhood removed from the justice system, using pre-emptive prevention, diversion and minimal intervention. All work minimises criminogenic stigma from contact with the system.

<sup>3</sup> System research: the multi-method ‘Edinburgh Study’ (McAra and McVie 2010) and survey-based ‘Swansea Bureau’ evaluation (Haines et al 2013) identified system contact and risk-based youth justice as criminogenic; Practice research: interviews/observations with practitioners identified risk assessment as marginalising children’s voices (Drake et al. 2014) and criminalising children (Creaney 2020); Process analyses: the risk-based ‘Scaled Approach’ framework has been significantly outperformed by Child First models (Haines and Case 2012); Critical reviews: evaluation of 39 meta-analyses/systematic reviews (Prins and Reich 2021) concluded that risk assessment claims to predictive accuracy are undermined by inappropriate statistics and inconsistent, overstated conclusions.

<sup>4</sup> The ‘Strategy Implementation Project’ (Case and Browning 2021b) consulted with a broad range of youth justice stakeholder groups and identified continued adherence to and privileging of risk management doctrine (e.g. within inspection criteria) as the key barrier to the successful/effective implementation of Child First in practice (see also Day 2022).

<sup>5</sup> For example, the Realist Synthesis review/evaluation methodology (Pawson and Tilley 1997) is specifically designed to unravel the complexities embedded in social programmes to understand how *context* shapes the *mechanisms* which result in interventions working and the *outcomes* they produce.

<sup>6</sup> Rather than more traditional, reductionist and static conceptualisations of context as observable features of interventions (space, place, people, things) that trigger or block interventions (Greenhalgh and Mazano 2021).

In this paper, we explore the use of the (quasi) realist EMMIE framework (see Thornton et al., 2019) for examining the content and rating the quality of evidence reviews in criminal justice, which guides users through a number of criteria that move beyond simply answering whether there is evidence that the intervention works or not. EMMIE explores intervention **E**ffects (impact on crime) but also whether the review presents any evidence of the heterogeneity of these effects, how they work (**M**echanisms or mediators), where they work (context or **M**oderators) and how interventions are **I**mplemented and their **E**conomic cost (hence 'EMMIE'). This framework was developed using adapted realist concepts to render them more acceptable to policy-makers and funders, serving as a bridge or mediator (read 'compromise') between the experimental, quantitative 'What Works' framework and more contextualised, qualitative realist approaches (Tilley 2018). Consequently, we selected EMMIE to explore the practicalities of using an adapted realist approach (Pawson, 2013; Emmel et al, 2018) with highly-rated (based on 'What Works' criteria) youth justice intervention evaluations. We conclude, however, that the paucity of contextual detail in these evaluations made it almost impossible to satisfy the information needed to assess effectiveness using the EMMIE framework - ultimately asserting that only Realist Synthesis can fill the decontextualised gaps in 'effectiveness' evaluations in youth justice and thus shed light on the contexts that shape the mechanisms through which youth justice interventions 'work'.

### **Decontextualised evaluation of youth crime prevention interventions: 'What Works' as the arbiter of 'effectiveness'**

Evaluation of the 'effectiveness' of youth crime prevention interventions in the Western world since the 1990s has been largely shaped by the 'What Works' framework, initiated and popularised by the (UK) text 'What works: Reducing reoffending' by James McGuire (1995) and consolidated in the eponymous (USA) study 'Preventing Crime: What Works, What Doesn't, What's Promising' (Sherman et al, 1998)<sup>7</sup>. The seven identified 'McGuire Principles'<sup>7</sup> (1995) for 'What Works' in reducing reoffending, subsequently underpinned much of the so-called 'effective practice for preventing youth offending in England and Wales.

Sherman et al's meta-evaluation of the 'effectiveness' of crime prevention programmes (USA-based) utilised 'scientifically recognized standards and methodologies', specifically the quantitative, experimental 'Scientific Methods Scale' (SMS – Cook and Campbell 1979), which confers methodological 'quality' on individual crime prevention evaluations - operationalised as *control* over extraneous variables, *measurement error* and *statistical power* to detect programme effects - examined across five analytical stages escalating in rigour and detail<sup>8</sup> (Sherman et al, 1998). Sherman et al privileged *youth* crime prevention interventions targeting 'factors that relate to juvenile crime including "risk factors" in the community, schools, and family' (ibid: v). These 'risk-focused interventions' were assigned 'What Works' effectiveness status if they: were 'reasonably certain' to prevent crime/reduce risk factors; had findings generalisable to similar settings in other places and

<sup>7</sup> *Risk classification* – level of service provided should match the level of risk assessed); *criminogenic needs* – intervention focus should address offending by alleviating needs that are criminogenic; *responsivity* - approaches that match worker and client learning styles work best; *community base* - programmes in the community fare better than those in institutions; *treatment modality* - effective interventions recognise that offenders experience multiple problems; *programme integrity* - effective interventions are well-designed and connects methods to aims; *dosage* – programmes need to be of sufficient intensity and duration..

<sup>8</sup>Correlation between a program and crime measure at one point in time; 2) Analysing measures of crime before and after the program - no comparable control condition; 3) Analysing measures of crime before and after the program in experimental and control conditions; 4) Analysing measures of crime before and after the program - multiple experimental and control conditions, controlling for extraneous variables (e.g. through matchings or statistical controls); Random assignment of units to program and control conditions - the Randomised Controlled Trial/RCT 'gold standard' of evaluation methodology

times; possessed at least two successful evaluations at SMS level three (or above<sup>9</sup>). Therefore, these criteria defined subsequent understanding of youth crime prevention intervention ‘success’ (reducing risk factors) and how success should be evaluated (via controlled trials).

As the 21st century began, the growing popularity of the McGuire Principles for practice and the ‘What Works’ framework for evaluating intervention effectiveness provided an ideal vehicle for risk management (e.g. early intervention targeting assessed risk factors), which became a popular model/strategy of youth justice across Anglophone countries, alongside the responsabilising and managerialist ‘new youth justice’ (Tompson, et al., 2020), which shaped the newly-constructed the Youth Justice System (YJS) of England and Wales (Case, 2022), with its primary aim to prevent youth re/offending. The ‘What Works’ promise of reducing recidivism through experiment-based treatment programmes ‘rooted in criminological knowledge’ and evaluated using ‘rigorous science’ (Cullen, 2005: 1) complemented neo-liberal and risk-based thinking, which was exponentially influencing the direction of youth justice policy from the late 1990s. The ‘What Works’ framework was immediately attractive to politicians because these reviews presented quick, apparently evidence-based, off-the-shelf solutions that prescribed ‘effectiveness’ grounded in risk management/reduction. Moreover, it provided professional practice with a modernising body of demonstrably ‘effective’ evidence-based crime prevention interventions that were, concurrently, defensible, transparent, replicable, coherent and easily applicable in practice. Accordingly, risk-focused prevention programmes became commonplace due to their emphasis across the Key Elements of Effective Practice guidance given to local youth offending teams (YOTs) by the Youth Justice Board<sup>10</sup> (YJB) to shape their ‘effective’ preventative responses to youth crime (Stephenson and Allen, 2013).

### **Promoting partiality: Reductionism and decontextualisation in the measurement of effectiveness**

The confluence of risk-led preventative interventions gaining prominence in youth justice practice and ‘What Works’ gaining prominence in evaluations of intervention ‘effectiveness’ engendered understandings of/responses to children’s offending that were inherently *reductionist* (e.g. over-simplified, narrow, restricted) and *decontextualised* (e.g. neglecting examination of broader contextual influences - socio-structural, relational, situational). Youth justice interventions considered effective through ‘What Works’ evaluation promoted a dual *partiality - bias* towards psychosocial, individualising, risk-led understandings of/responses to youth offending and an *incompleteness* in terms of understanding youth offending (see Sutton et al, 2022). The risk-based preventative interventions favoured by the SMS-led ‘What Works’ evaluation framework (when applied to youth justice), therefore, generated a narrow evidence-base through imposing rigid, experimental measures of ‘quality’ and ‘effectiveness’, often based on explicit randomisation into control and treatment groups. Indeed, Sherman et al (1998) acknowledged that US Congress specifically required them to provide a ‘comprehensive’ evaluation of effectiveness, yet giving special emphasis to “(a) reductions in delinquency, juvenile crime, youth gang activity, youth substance abuse, and other high-risk factors; (b) reductions in the risk factors in the community, schools, and family environments that contribute to juvenile violence; and (c) increases in the protective factors that reduce the likelihood of delinquency and criminal behavior”. Synonymising evidence quality and effectiveness measures with experimental research offered an attractively prescriptive (managerialist) and modernising ‘evidence-based’ approach to neo-liberal governance, yet the consequent reductionism and decontextualisation seriously

<sup>9</sup> Programmes were labelled as ‘what’s promising’ if the level of certainty regarding reduction of crime or risk factors was too low to support generalisable conclusions, but there was some empirical basis for predicting that further research could support such conclusions, in addition to at least one successful evaluation at level three. Sherman et al. (1998) classified ‘what doesn’t work’ as programmes that they were reasonably certain failed to prevent crime or reduce risk factors (or rather, cannot be measured as reducing crime and reducing risk according to the SMS).

<sup>10</sup> A non-departmental public body leading youth justice strategy through advising government on policy development and advising youth offending teams and how to most effectively animate policy and practice.

1  
2 questions whether the approach provides youth justice stakeholders with a comprehensive and valid  
3 evidence-base to guide ‘effective’ preventative interventions.  
4

5 More recent moves towards contextualising youth justice intervention effectiveness have struggled to  
6 shed their reductionist shackles. For example, the YJB’s ‘Theory of Change’ guidance for  
7 understanding how ‘effective’ preventative interventions might ‘work’ (YJB, 2016) introduced a new  
8 focus on ‘context’, albeit operationalised narrowly as observable static ‘features’ of situations or  
9 populations, rather than dynamic (changeable) and relational features shaping the mechanisms through  
10 which interventions work (Greenhalgh and Manzano, 2021). Consequently, whilst context has been  
11 acknowledged as crucial to effectively evaluating mechanisms or causality in intervention success (see  
12 also Adler et al, 2016 for the Ministry of Justice (MOJ)), the construction and operationalisation of  
13 evaluation embedded within YJB and MOJ guidance remains reductionist and, ironically, largely  
14 decontextualised (Sutton et al, 2022). Even the increasingly popular ‘What Works’ Centres (e.g. the  
15 UK-based What Works Centre for Crime Reduction) have acknowledged that their traditionally-  
16 accepted ‘robust’ methods for evaluating the ‘effectiveness’ of youth crime prevention interventions  
17 should now be accompanied by a more systematic and contextualised analysis of what is working,  
18 where and why (Tompson et al, 2020), but have been slow to take up this challenge.  
19  
20  
21

### 22 **Contextualising evaluations of youth justice intervention: Realist methods**

23 Realist researchers have become increasingly interested in the application of alternative,  
24 methodologies offering more nuanced and coordinated evaluation of the ‘effectiveness’ of preventative  
25 interventions in the youth justice field. Realist approaches focus on *how* and *why* interventions work,  
26 for whom and in what contexts – expressed as ‘Context-Mechanisms-Outcomes’ configurations and  
27 are central to both realist evaluation (Pawson and Tilley, 1997) and realist synthesis (cf. Pawson,  
28 2006). Whereas experimental approaches to evaluation tend to be driven by the collection of discrete  
29 data relating to dependent and independent variables, realist approaches are driven by the theories that  
30 underlie interventions. Realists foreground a ‘generative’ understanding of causation (Pawson, 2006),  
31 which seeks to understand the *mechanisms* through which youth justice interventions can help children  
32 to desist from offending – mechanisms typically neglected in the experimentalist ‘What Works’  
33 framework. A key benefit of taking a realist approach to the evaluation of interventions, therefore, is  
34 that it considers the elements that underlie mechanisms of change, exploring context as ‘the relational  
35 and dynamic features that shape the mechanisms through which the intervention works’ (Greenhalgh  
36 and Manzano, 2021: 1). As realism is philosophically-situated between positivism and constructivism,  
37 it facilitates broader, more in-depth analyses of the causal mechanisms shaping intervention  
38 effectiveness, moving beyond the partial explanatory utility of ‘successionist<sup>11</sup>’ positivist models. In  
39 contrast, a ‘generative’ understanding of causation underpins the realist approach to synthesis. This  
40 seeks to explore how contextual features shape the mechanisms through which interventions produce  
41 their outcomes. As mechanisms do not often lend themselves to measurement, unlike variables  
42 (Pawson, 2006a), generative causation seeks to understand processes which describe the actions that  
43 lead to ‘demi-regularities’ are in outcomes (i.e. the relatively stable occurrence of an observable event,  
44 which are dependent on the choice-making capacity of individuals and groups (Lawson 1997). The  
45 promise for evaluation in youth justice is that, rather than summing evidence from experimental and  
46 quasi-experimental studies to isolate the average effect size of an intervention, realists seek to explore  
47 how contextual features shape the mechanisms through which interventions produce their outcomes  
48 by integrating both quantitative and qualitative evidence<sup>12</sup>. This fuller, more in-depth understanding  
49  
50  
51  
52  
53

54  
55 <sup>11</sup> Successionism examines the causal relationship between dependent and independent variables, such that observing these  
56 causal associations between (in/dependent) variables becomes a closed system that rules out alternative explanations and  
57 experimental approaches become the gold standard.

58 <sup>12</sup> For example, Petrucci (2021) illustrates that an intervention’s success or failure can involve complex interactions  
59 within varied contexts, so the context-mechanism-outcome hypotheses of realist evaluation can incorporate such  
60 complexity and allow for process and outcome patterns to be measured simultaneously, enriching the conceptualisation  
and measurement of intervention.

could provide policymakers and practitioners with a stronger evidence-base, which can help them better support children and to prevent offending because stakeholders are encouraged and enabled to view these children as individuals and to design more nuanced preventative interventions as a result.

### **Applying EMMIE to better understand and contextualise preventative interventions**

As discussed in the Introduction, EMMIE<sup>13</sup> provides a framework for examining the extent to which reviews provide information not just about the effect of interventions, but also the causal mechanisms through which they work, the contextual features that shape the way interventions work, their impact and practical implementation issues for policymakers. Importantly, EMMIE also seeks to ascertain whether reviews provide evidence of value for money. Employing this nuanced evaluation methodology alongside a realist review this can enable researchers to: identify a broader range of contextualised criminal mechanisms; reveal intervention needs; and explore the contextual features (e.g. relational, dynamic, socio-structural) that shape the mechanisms through which interventions work (Greenhalgh and Manzano, 2021).

### **Methodology: The EMMIE review of reviews**

As part of a Nuffield Foundation-funded project entitled ‘Understanding criminogenic influences on youth offending: Context, mechanism, outcomes’ (cf. Sutton et al, 2022), we sought to examine how criminogenic and contextual influences are constructed (and neglected) in youth crime prevention interventions and thus how they could be targeted to produce more effective interventions. Accordingly, we proposed an EMMIE-informed ‘review of reviews’, predominantly to evaluate intervention effects and economic costs (the ‘E’ elements of EMMIE), whilst also enabling exploration of the MMI elements (moderators, mechanisms and implementation). Therefore, EMMIE supported evaluation of the content and quality of a sample of existing youth justice reviews of the effectiveness of youth justice interventions (meta-analyses/meta-evaluations, not stand-alone interventions), identified as part of a ‘review of reviews’ (ROR). This afforded a more efficient means of accessing a broad cross-section of intervention literature over our limited project timeframe. The dual aim for the ROR exercise was to test the applicability and utility of EMMIE for analysing and evaluating reviews of the effectiveness of preventative interventions in the youth justice sphere. We also wanted to examine evidence of varying offending outcome patterns across effectiveness reviews, which could identify whether evaluations share contextual features or intervention resources driving positive impacts on intended outcomes. We extended and refined the approach adopted in a comprehensive EMMIE review by Tompson et al (2020), which examined the quality of systematic review evidence of crime reduction intervention effectiveness. Tompson et al (2020) used EMMIE to examine the nature of evidence of effectiveness contained within 70 systematic reviews of single crime reduction initiatives over the period 1975-2015. We mirrored this approach by focusing on single intervention reviews (contained in the Campbell Library) and adopting an equivalent search protocol. We extended their approach by incorporating research not captured by the Tompson et al review (e.g. evaluations published after their 2015 review cut-off) and refined their approach by focusing our review on *youth* justice interventions.

### **Sampling: Identifying relevant reviews**

As a key member of the Tompson et al review team sits on our project advisory board, we were able to solicit expert advice on our approach to identifying review literature. He advised that the searching aspect of their review would be problematic due to the limited time frame allocated to the exercise (September-December 2021) and the sizable workload associated with screening up to 10,000 abstracts. Consequently, a more restricted, selectively-focused search strategy was required. Accordingly, we were supported in identifying the Youth Endowment Fund (YEF) Evidence and Gap Map/EGM (YEF, 2021) as an appropriate sampling frame. The EGM constitutes a large scale,

---

<sup>13</sup> A framework exploring intervention **E**ffects (impact on crime), how interventions work (**M**echanisms or mediators), where they work (context or **M**oderators) and how they are **I**mplemented and their **E**conomic cost (hence ‘EMMIE’).

1  
2 structured resource for identifying evidence on interventions for 'preventing children getting involved  
3 in violence'. It claims to illustrate the volume and the different levels of quality of evidence of  
4 effectiveness by rating studies against specified criteria in their quality assessment tool, which was  
5 aligned to Campbell Collaboration methodology. Although restricted in outcome focus (youth violence  
6 only, not youth crime), the EGM still identified over 2000 studies (from searches in 2020), thus  
7 providing the largest, most up-to-date resource of youth violence prevention studies in the world at  
8 that time (YEF, 2021). Our decision to utilise the EGM as the basis of our sampling frame was  
9 informed by:

- 10  
11  
12 • **Challenges:** collating a large body of literature poses extensive time, research and intellectual  
13 challenges to the research team, which could be mediated through adopting and adapting the  
14 Tompson et al (2020) inclusion-exclusion criteria (see below);
- 15  
16 • **Screening and inclusion criteria:** screening against inclusion criteria for studies entering the  
17 EGM removed reviews that would not be relevant to us because they did not relate to violent  
18 offences. The EGM inclusion criteria were broader than ours because they included primary  
19 research publications as well as systematic reviews. This meant that the EGM should contain most  
20 of the studies relevant to our scope, as a subset of their map.
- 21  
22 • **Transparency:** the original search methods were reported fully. This allowed them to be checked  
23 for rigour, using the Peer Review of Electronic Search Strategies/PRESS checklist (McGowan, et  
24 al., 2016) and Campbell Collaboration guidance for systematic review search methods (Kugley, et  
25 al., 2016), and enabled their application in updated searches that could also capture reviews  
26 published since the EGM conducted their searches;
- 27  
28 • **Rigour and sensitivity:** prioritising published search strategies and source lists facilitated our  
29 rigorous and sensitive search strategy, which was likely to identify the most relevant reviews for  
30 updated searches;
- 31  
32 • **Incentives:** we remained attentive to the moral and economic incentives for re-using previous  
33 rigorous research, rather than incurring research waste by duplicating the work of the EGM in  
34 identifying and selecting the same relevant studies.

### 37 38 **Search methods**

39 To identify reviews relevant to the scope of our ROR exercise, we gathered systematic reviews of  
40 youth violence preventative interventions from the EGM and conducted update searches of databases  
41 and websites following advice from the Information Specialists involved in the EGM creation. The  
42 update searches of the databases and resources used for the EGM ensured the team identified the latest  
43 reviews. Records of systematic reviews were extracted from the EGM. Approximately 10% of the  
44 systematic review records were tagged with our key outcome 'offences'. After checking, we  
45 discovered that some records without the 'offences' tag were relevant to our review, so we decided to  
46 extract and screen all systematic review records. We replicated EGM searches as closely as possible  
47 in terms of strategies, databases and website resources, searched. Searches undertaken in the databases  
48 included search concepts for: children and adolescents, behaviour, crime prevention, social support  
49 interventions, gangs and drug use involvement, family interventions, school and community support  
50 and criminal justice. We added a further search concept for 'systematic reviews' to remove non-  
51 reviews from our results. To capture records added to sources since the EGM conducted their searches,  
52 we used entry date search fields (where available) to limit searches to only records added since June  
53 2020. The EMG and update searches were further supplemented by a set of potentially relevant reviews  
54 sent from the Campbell Collaboration Information Specialists, which linked to their ongoing EGM  
55 work. The results of the database searches were stored in an online Rayyan database ([www.rayyan.ai](http://www.rayyan.ai)),  
56 so that colleagues could access this information to undertake inclusion and exclusion activities  
57 remotely.  
58  
59  
60

### Inclusion-exclusion criteria for identifying reviews

We developed and applied a set of inclusion and exclusion criteria based on a Population, Intervention, Comparator and Outcome (PICO) framework (Richardson et al, 1995), an established framework for formulating review questions about the effectiveness of interventions (Cochrane Collaboration, 2011):

- **Population:** children aged 10-17 identified as having offended;
- **Intervention):** single preventative interventions;
- **Comparator:** not applicable
- **Outcome:** effectiveness of interventions measured by rates of re-offending

The 'P' was guided by the scope of our realist review (i.e. official responses to offending by young people); the 'I' and 'O' on by our decision to mirror, adopt and adapt the methodology of Tompson et al. We decided to filter our results to focus on prominent reviews (indicated by their eligibility for publication in Campbell Systematic Reviews or the Cochrane Database of Systematic Reviews) stemmed from the 'gold standard' moniker attached to 'What Works' style reviews. The Campbell Collaboration claim to produce unbiased, accurate summaries of empirical social science evidence using systematic review methods, meta-analysis, and related approaches (Littell and White, 2018). Cochrane Reviews base findings on health research results which meet specific quality criteria and apply methods intended to reduce the impact of bias across different parts of the review process (Cochrane Library, 2022). Therefore, we added the Campbell-Cochrane selection criteria because both types of systematic review follow a similar approach to those prioritised by 'What Works' reviews. We believed that these reviews were more likely to include relevant research most reflective of the common forms of youth justice intervention 'effectiveness' and the type of (narrow, reductionist, decontextualised) 'evidence' privileged by youth justice stakeholders trust to inform policy, practice and decision-making.

The literature search identified six reviews:

1. **Cognitive-behavioural therapy (CBT) interventions for young people (Sneddon et. al, 2020):** Examined whether CBT was better at reducing adolescent harmful sexual behaviour than no treatment or treatment as usual.
2. **Drug Courts' effects on criminal offending for juveniles and adults (Mitchell et. al, 2012):** Reviewed the effects of Drug Courts on general reoffending for juveniles using a meta-analysis of 154 independent evaluations of drug court programs.
3. **Formal system processing of juveniles: Effects on delinquency (Petrosino et. al, 2010):** Considered how formally processing children through justice systems effected future delinquency through the meta-analysis of 29 controlled trials.
4. **Restorative justice conferencing for reducing recidivism in young offenders (Livingstone et. al, 2013):** Meta-analysis of restorative justice conferencing impacts on children's propensity to reoffend, alongside secondary measures including remorse, recognition of wrongdoing, satisfaction and self-perception.
5. **Police-initiated diversion for youth to prevent future delinquent behavior (Wilson et. al, 2018):** Reviewed the impact of police-initiated diversion on rates of juvenile reoffending compared to standard processing in the justice system.
6. **Multisystemic therapy for social, emotional, and behavioral problems in youth (Littell et. al, 2005):** Reviewed the effectiveness of licenced multisystemic therapy programmes with children aged 10-17 classed as delinquent or from other vulnerable groups of children.

### Results: EMMIE review of reviews

We examined each review systematically against the EMMIE components:

- Effectiveness: Evidence criteria includes reporting of effect size, moderator analysis and the measurement of unanticipated effects. Quality scores in this category replicate many of the traditional quality criteria for assessing trials in terms of risk of bias;
- Mediators/Mechanisms: Evidence criteria includes maps of mechanisms, priori mediator or mechanism-based moderator analysis, post hoc or mechanism-based moderator analysis and assessment of most likely mechanisms and any contextual conditions. Quality scores relate to the level of consideration given to program theories in the evidence;
- Moderators/Context: Evidence criteria includes statements qualifying contextual variation, a priori or post hoc context-based moderator or subgroup analysis; quality scores relate to the level of consideration of moderators or contexts.
- Implementation: Evidence criteria focusses on the statement of key components necessary for implementation or replication; quality scores relate to implementation issues;
- Economics: Evidence criteria includes quantification of inputs, outputs, intensity and estimation of costs and benefits; quality scores relate to the level of detail in these respects.

The limitations of the content in these reviews became immediately apparent in the first review and were reflected in each subsequent review. Whilst the reviews ostensibly cohered around ‘gold standard’ effectiveness evidence, their relative neglect of non-RCT research was lamented by several review authors, notably in relation to their consequent inability to fully explore features of EMMIE beyond ‘Effectiveness’ due to a lack of detailed information (i.e. methodological limitations cf. the Sneddon, Mitchell and Livingstone and Petrosino studies).

A summary of findings against each EMMIE element is provided in Table 1 and synthesised below it.

#### TABLE 1 HERE

- **Effectiveness:** Effect sizes in relation to reoffending rates were common measures of effectiveness in all the selected studies, with authors either conducting meta-analyses or simply presenting effect size for each included evaluation, but they rarely provided a narrative overview. It was difficult to gain an understanding of whether and how the strength of effect varied across studies and associated contextual details across these youth justice intervention reviews was a notable weakness. Whilst meta-analyses of RCT/quasi-experimental studies appear to be the dominant approach to reviewing intervention effectiveness in our sample, it is evident that this approach neglects important variations in effect size.
- **Mechanisms:** Whilst there was some consideration of Mechanisms in the sample reviews and some useful findings highlighted, such as the impact of court demands and intervention duration (Mitchell et al, 2013) and diversion being more effective when children view it as onerous (Petrosino et al, 2010). Some review authors (e.g. Sneddon et al, 2020; Mitchell et al, 2012) suggested that they would have liked to have seen detail examining the *mechanisms of change*, but there was often insufficient information in evaluations to facilitate this (Livingstone et al, 2013). This reflects a growing appetite for ‘robust’ evaluations of ‘effectiveness’ of youth crime prevention interventions coupled with a systematic analysis of what is working, where and why. However, our review confirms that this information, as evident in Tompson et al’s (2020) review, continues to be sparse and when it is provided it is perceived as poor quality and often lacks the

necessary detail to facilitate understanding of the criminogenic or cessation mechanisms that can occur from preventative intervention.

- **Moderators:** The quality of Moderators evidence was better in respect of the two reviews of diversion (Petrosino et al, 2010; Wilson et al, 2018) and the review of the drug courts (Mitchell et al, 2012), than it was in the other three reviews (see table). For example, reviews by Sneddon et al (2020) and Littell et al (2005) contain no moderator information because of the limited detail in their included research. Where moderator analyses were conducted there were concerns expressed about heterogeneity in the sample (Livingstone et al, 2013) and a lack of statistically significant differences between the intervention and control groups in the included studies suggested these differences were not meaningful (Wilson et al, 2018). This contrasts to the findings of Tompson et al, (2020) who found that moderator analyses were good overall in their sample of preventative intervention reviews.
- **Implementation:** We found that reviews either failed to consider Implementation (Sneddon et al, 2020; Littell et al, 2005) or were unable to consider this due to a lack of detail in the selected research (e.g. Livingstone et al, 2013; Wilson et al. 2018). Most of the discussion of potential implementation issues in this sample of reviews was unsubstantiated, for example, Littell et al, (2005) outline that Multi-systemic therapy outcomes may be limited by the short-term nature of individual and family-focused interventions, regardless of design or intention, although there was no evidence in their sample of evaluations to support this.
- **Economic cost:** The ROR revealed that detail on Economic cost was completely lacking, as also highlighted as a concern in the Tompson et al (2020) review. Only one review estimated costs of delivering the intervention (Littell et al, 2005) and none of the reviews were able to conduct cost-benefit analyses due to a lack of evidence, so any discussion of potential cost-benefits was unsubstantiated.

### Discussion: The promise of EMMIE?

The EMMIE framework has become a key component of the systematic review/evaluation landscape in terms of understanding the (adult) crime prevention evidence-base, yet we believe that our study is the first attempt to apply this framework to *youth* justice interventions designed to prevent offending. The clear promise of EMMIE for the evaluation of the effectiveness of preventative interventions in the youth justice field is that it provides a framework to examine the extent to which reviews not only examine intervention *effects*, but also the causal *mechanisms* through which they work, the contextual features that shape the way interventions work and their impact and the practical implementation issues for policymakers. Indeed, the EMMIE framework is filtering into ‘What Works’ networks<sup>14</sup>—demonstrating that intervention programme design and implementation can be better informed by realist methods enabling consideration of contextual features and causal mechanisms neglected by the ‘What Works’ approach. These include how the implementation of an intervention can shape its effectiveness and its impact on a child, as well as considering individual, family, social, economic and other features. The EMMIE framework encourages researchers to engage in further kinds of interrogation when evaluating (youth justice) interventions, which can extend understandings of the efficacy of different responses to youth crime by considering contextual differences and causal mechanisms.

<sup>14</sup> e.g. Providing the methodological foundation for the UK ‘What Works’ Centre for Crime Reduction, hosted by the College of Policing (<https://whatworks.college.police.uk/About/Pages/default.aspx> accessed on 06 June 2021).

1  
2  
3 However, the reportage of previous studies in the reviews that we included suggests that the requisite  
4 evidence for EMMIE features is not always included or (at least) prioritised in youth justice  
5 intervention reviews or the standalone studies they incorporate). In particular, the strength of effects  
6 reported across the reviews was weak, which made it difficult to evaluate the effectiveness of  
7 interventions and how/why they may work with children. Furthermore, there is little appreciation of  
8 the influence of context in these analyses to enable consideration of whether and how the effectiveness  
9 of these interventions varies with children with different background characteristics. However, by  
10 deliberately reducing the scope of research to experimental and RCT methods, these reviews appeared  
11 to actively restrict the detail the researchers would have liked to have seen to examine why such effects  
12 were evident. This was particularly the case in respect of understanding how contextual features might  
13 shape the intervention effects and also in considering different elements of implementation and their  
14 impact and in terms of *cost*-effectiveness. For example, whilst every review considered the strength of  
15 intervention effects, there was little detail pertaining to *how* any of these interventions might ‘work’  
16 in preventing offending and how different contextual features might limit or maximise preventative  
17 effects. Therefore, our ability to examine the EMMIE elements of chosen studies was restricted by the  
18 limited information contained in the ‘What Works’ style reviews of youth justice interventions we  
19 selected, despite this review data being heralded as the evidential ‘gold standard’ in determining  
20 intervention effectiveness and quality.  
21  
22  
23  
24

25 It would appear, therefore, that application of the EMMIE framework cannot immediately ‘fix’ the  
26 inherent limitations in the dominant youth justice evidence-base, but it can certainly help to expose the  
27 gaps in what is included in these reviews. Furthermore, it is evident that ‘What Works’ reviews in  
28 youth justice are often only really designed to answer limited elements of the first ‘E’ in EMMIE –  
29 namely effect sizes. However, it is highly probable that incorporating reviews adopting a more  
30 inclusive sampling approach and less restricted methodological selection criteria might yield more  
31 contextual detail than those with a ‘What Works’ experimental focus in reviewing evidence. For  
32 example, qualitative and practitioner research, which tend to be factored out of ‘gold standard’ reviews,  
33 might be used in combination with experimental research to help in bridging these gaps in  
34 understanding (Johnson et al, 2015). In particular, crucial information required for understanding the  
35 mechanisms at play in interventions and the contextual features promoting effectiveness might be  
36 gleaned from so called lower quality, non-experimental evaluations of evidence. This is indicated  
37 strongly by Mitchell et al’s (2012: 24) assertion (regarding drug courts) that a predisposition for  
38 experimental rigour can severely limit demonstration of effectiveness in evaluations of preventative  
39 youth justice interventions. Accordingly, the valorisation of experimental research as *the* gold standard  
40 can be seen to be restrictive in understanding factors at play in youth justice interventions which make  
41 them successful or unsuccessful in promoting recidivism. Therefore, whilst EMMIE provides a  
42 potentially welcome addition to the youth justice review infrastructure, it is currently hamstrung by  
43 the reporting of findings in hegemonic ‘What Works’ reviews - a framework that will not support our  
44 quest for further contextual detail to better understand preventative intervention effectiveness for  
45 children who offend. By piloting use of this framework, we have highlighted that the dominant body  
46 of evidence in youth justice produces challenges for EMMIE, mainly because this body of evidence  
47 does not present or indeed seek enough of these types of data to make it feasible for use. Notably, the  
48 lack of consideration of the context of intervention effectiveness in this sample of reviews, implies that  
49 the quantifying, aggregating processes of typical ‘What Works’ interventions smooths over/washes  
50 away individual differences and that EMMIE struggles to reveal why youth justice interventions work  
51 for some children and not for others, or why they are only effective in certain situations and  
52 circumstances. Whilst EMMIE (as a purported middle-ground between positivism and realism - Tilley,  
53 2018) in theory offers potential insight into evaluating the impact of contextual differences on youth  
54 justice interventions in the quantitative, experimental evidence-base, in practice this evidence-base  
55  
56  
57  
58  
59  
60

appears bereft of essential information regarding the context and mechanisms of change within interventions.

### The promise of Realist Synthesis?

As illustrated by our ROR, syntheses can only work when the requisite evidence is available. Accordingly, the ROR suggests that when evaluating reviews of youth justice interventions, the Mechanism-Moderator-Implementation components of EMMIE may be better addressed through Realist Synthesis review. Realist Synthesis an established review methodology (Pawson and Tilley 1997) specifically designed to disentangle the heterogeneity and complexity of social programmes , aiming to go beyond an estimation of ‘What Works’ to understand how *context* shapes the *mechanisms* through which interventions work and the *outcomes* that interventions produce. Realist Synthesis offers an iterative process of identifying, testing and refining these programme theories to build explanations about how and in what circumstances these interventions work and why through a synthesis of literature using a wide range of study designs. The advantage of Realist Synthesis here, therefore, is that the approach makes use of a much broader range of study designs than the ‘What Works’ interventions prioritised and privileged by our selected reviews evaluated through EMMIE. Crucially, unlike many other review strategies, Realist Synthesis is theory-driven (Blamey and Mackenzie 2006), prioritising intervention *theory* and marshalling a broader spectrum of evidence to ascertain how programmes work, for whom and in what circumstances. Realist approaches move beyond the explanatory ‘black box’ of successionist models to offer a generative understanding of causation. The promise for evaluation in youth justice is that, rather than summing evidence from quasi/experimental studies to isolate the average effect sizes of interventions, realist approaches integrate both quantitative and qualitative evidence to explore how contextual features shape the mechanisms through which interventions produce their outcomes and how contextual features may impact upon intervention implementation. Consequently, the reductionist and decontextualising limitations of the ‘What Works’ approach and the responsabilising and adulterising excesses of the interventions approaches it facilitates and validates could be addressed through the use of Realist Synthesis to identify and synthesise a broader evidence-base for better understanding and addressing (in policy and practice) causal mechanisms and the role of context in youth offending and its prevention (Sutton et al, 2022). Realist Synthesis starts by prioritising the underlying programme theories/theories of change as the basis for evaluating intervention research literature (and is explicitly concerned with understanding how contextual features (e.g. the Child First contextual priority of positive relationship-based practice) shape the mechanisms through which interventions work (e.g. positive relationships encouraging the development or enhancement of prosocial identity) to produce un/intended outcomes. As such, it is equipped to synthesise the sorts of evidence and answer the sorts of questions that fit the MMI components of the EMMIE framework, which do not appear to be available in dominant methodologies for evaluating youth justice interventions. This evidence provides an important foundation for examining how context shapes both the implementation of the intervention and how people respond to the resources that are offered (or taken away) by interventions, which can shed light on what mechanisms are assumed to be at play (Pawson, 2013, 2006b). Furthermore, the emerging ‘Child First’ environment for youth justice practice is far more context-sensitive than its adult-centric, decontextualised risk management-led predecessor, with its component tenets having the clear potential to explicitly addressing contextual issues at multiple levels of the social system: macro (e.g. see children as children by acknowledging structural barriers), meso (e.g. develop prosocial identity and collaboration with children built on supportive relationships) and micro (e.g. develop prosocial identity to promote children’s strengths and potential). Consequently, the inherently contextualised nature of Child First as the guiding principle for youth justice is-can be promotive of contextualised understandings of and responses to children who offend (e.g. through context-sensitive, child-centric preventative interventions) and should therefore be compatible (even synergistic) with the contextualised evaluation approach afforded by Realist Synthesis.

1  
2  
3  
4 **Conclusion: In support of Realist Synthesis evaluation of intervention ‘effectiveness’**

5 As a result of this study, we advocate for the adoption of the Realist Synthesis evaluation methodology  
6 in the youth justice field to enable more holistic and valid understandings of how criminogenic  
7 influences and programme theories have been conceptualised in youth justice interventions. Compared  
8 to the hegemonic ‘What Works’ evaluation framework and even compared to the more synergistic  
9 EMMIE approach, Realist Synthesis offers the sector far more capacity and utility when seeking to  
10 identify contextualised explanations of youth offending. Realist Synthesis examines both the causal  
11 mechanisms through which interventions work (or do not work) to produce their intended and  
12 unintended effects and it also interrogates the vital contextual influences on the mechanisms of change  
13 that underpin intervention effectiveness. Adopting this approach would enable stakeholders to identify  
14 where there are implementation issues in youth justice and/or fundamental flaws in the implicit  
15 assumptions about causality and how interventions work. Therefore, Realist Synthesis evaluation  
16 could broaden the explanatory scope of the existing youth justice evidence-base, so the understanding  
17 of which programmes work could be further expanded, and consideration could also be given to whom  
18 programmes work for, in what circumstances and why. Moreover, applying Realist Synthesis would  
19 facilitate more detailed reviews of the ‘effectiveness’ of preventative interventions in youth justice that  
20 are more sensitive to the complexities of children’s differing experiences, engagements and local  
21 circumstances. We believe that this offers a platform for synthesising the contextual elements of youth  
22 offending and youth justice intervention programmes that are often neglected in contemporary  
23 evidence-based practice.  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- Adler, J. R., Edwards, S. K., Scally, M., Gill, D., Puniskis, M. J., Gekoski, A., & Horvath, M. A. (2016). *What works in managing young people who offend?* London: Ministry of Justice.
- Astbury, B. (2018). Making Claims Using Realist Methods. In N. Emmel, J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing Realist Research* (pp. 59-78). London: Sage Publications.
- Blamey, A., & Mackenzie, M. (2007). Theories of change and realistic evaluation: peas in a pod or apples and oranges? *Evaluation*, 13(4), 439-455.
- Case, S.P. (2021) Challenging the reductionism of 'evidence-based' youth justice. *Sustainability*. Online first.
- [Case, S.P. and Haines, K.R. \(2009\) \*Understanding youth offending: Risk factor research, policy and practice\*. Cullompton: Willan.](#)
- [Case, S.P. and Hazel, N. \(in press\) \*Child First: Developing a new youth justice system\*. London: Palgrave.](#)
- [Case, S.P. and Browning, A. \(2021a\) \*Child First: The evidence-base\*. <https://www.lboro.ac.uk/subjects/social-policy-studies/research/child-first-justice/#:~:text=Child%20First%20Justice%20is%20a,rather%20than%20as%20offenders>.](https://www.lboro.ac.uk/subjects/social-policy-studies/research/child-first-justice/#:~:text=Child%20First%20Justice%20is%20a,rather%20than%20as%20offenders)
- [Case, S.P. and Browning, A. \(2021b\) \*The Child First Strategy Implementation Project\*. <https://www.lboro.ac.uk/subjects/social-policy-studies/research/child-first-justice/#:~:text=Child%20First%20Justice%20is%20a,rather%20than%20as%20offenders>.](https://www.lboro.ac.uk/subjects/social-policy-studies/research/child-first-justice/#:~:text=Child%20First%20Justice%20is%20a,rather%20than%20as%20offenders)
- Cochrane Library. (2022, January 31). *About Cochrane Reviews*. Retrieved from [www.cochranelibrary.com: https://www.cochranelibrary.com/about/about-cochrane-reviews](https://www.cochranelibrary.com/about/about-cochrane-reviews)
- Cook, T., & Campbell, D. (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand-McNally.
- Cullen, F. (2005). The Twelve People Who Saved Rehabilitaton: How the Science of Criminology Made a Difference. *Criminology*, 43(1), 1-42.
- [Day, A. \(2022\) It's a Hard Balance to Find': The Perspectives of Youth Justice Practitioners in England on the Place of 'Risk' in an Emerging 'Child-First' World. \*Youth Justice\*, Online first.](#)
- Emmel, N., Greenhalgh, J., Manzano, A., Monaghan, M., & Dalkin, S. (2018). *Doing Realist Research*. (N. Emmel, J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin, Eds.) London: Sage Publications.
- Farrington, D. (2000). Explaining and Preventing Crime: The Globalisation of Knowledge. *Criminology*, 38(1), 1-24. doi:10.1111/j.1745-9125.2000.tb00881.x

- 1  
2  
3 Fullerton, D., Bamber, J., & Redmond, S. (2021). *Developing Effective Relationships Between Youth*  
4 *Justice Workers and Young People: A Synthesis of the Evidence, REPPP Review*. Limerick:  
5 University of Limerick.  
6  
7  
8 Greenhalgh, J., & Manzano, A. (2021). Understanding ‘Context’ in Realist Evaluation and Synthesis.  
9 *International Journal of Social Research Methodology*, 1-13.  
10 doi:10.1080/13645579.2021.1918484  
11  
12  
13 Guyatt, G. H., Haynes, R. B., Jaeschke, R. Z., Cook, D. J., Green, L., Naylor, C. D., . . . Richardson,  
14 W. S. (2000). Users Guide to the Medical Literature XXV. Evidence-Based Medicine:  
15 Principles for Applying the Users Guide to Patient Care. *JAMA*, 284(10).  
16  
17  
18 Hendrick, H. (2015). Histories of Youth Crime and Youth Justice. In B. Goldson, & J. Muncie (Eds.),  
19 *Youth Crime & Justice* (pp. 3-16). London: Sage Publications.  
20  
21  
22 Johnson, S., Tilley, N., & Bowers, K. (2015). Introducing EMMIE: an evidence rating scale to  
23 encourage mixed-method crime prevention synthesis reviews. *Journal of Experimental*  
24 *Criminology*, 459-473.  
25  
26 Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A., Hammerstrøm, K., & Sathe, N. (2016).  
27 Searching for studies: A guide to information retrieval for Campbell. *Campbell Systematic*  
28 *Reviews*.  
29  
30  
31 Lipsey, M., & Wilson, D. (1993). The efficacy of psychological, educational, and behavioral treatment:  
32 confirmation from meta-analysis. *American Psychologist*, 48(12), 1181–1209.  
33  
34  
35 Littell, J., & White, H. (2018). The Campbell Collaboration: Providing Better Evidence for a Better  
36 World. *Research on Social Work Practice*, 28(1), 6-12. doi:10.1177/1049731517703748  
37  
38 Littell, J., Popa, M., & Forsythe, B. (2005). Multisystemic therapy for social, emotional, and behavioral  
39 problems in youth aged 10-17. *Campbell Systematic Reviews*.  
40  
41  
42 Livingstone, N., Macdonald, G., & Carr, N. (2013). Restorative justice conferencing for reducing  
43 recidivism in young offenders (aged 7 to 21). *Cochrane Database of Systematic Reviews*.  
44  
45  
46 McGowan, J., Sampson, M., Salzwedel, D., Cogo, E., Foerster, V., & Lefebvre, C. (2016). PRESS  
47 peer review of electronic search strategies: 2015 guideline statement. *Journal of clinical*  
48 *epidemiology*(75), 40-46.  
49  
50  
51 McGuire, J. (Ed.). (1995). *What works: Reducing reoffending: Guidelines from research and*  
52 *practice*. John Wiley & Sons  
53  
54 Mitchell, O., Wilson, D., Eggers, A., & MacKenzie, D. (2012). Drug Courts’ effects on criminal  
55 offending for juveniles and adults. *Campbell Systematic Reviews*.  
56  
57  
58 Muncie, J. (2011). Illusions of Difference: Comparative Youth Justice in the Devolved United  
59 Kingdom. *British Journal of Criminology*, 51, 40-57.  
60  
61 Muncie, J. (2014). *Youth and Crime: A Critical Introduction*. London: Sage.

- 1  
2  
3 Noble, J., & Hodgson, L. (2016). *Theory of Change Guidance*. London: YJB.
- 4  
5 Pawson, R. (2002). Evidence Based Policy: in Search of a Method. *Evaluation*, 157-81.
- 6  
7 Pawson, R. (2002). Evidence-based policy: The promise of 'Realist Synthesis'. *Evaluation*, 8, 340–  
8 358.
- 9  
10  
11 Pawson, R. (2006a). *Evidence-Based Policy: A Realist Perspective*. London: Sage.
- 12  
13 Pawson, R. (2006b). Digging for Nuggets: How 'Bad' Research Can Yield 'Good' Evidence.  
14 *International Journal of Social Research Methodology*, 127-142.
- 15  
16 Pawson, R. (2008). *Causality for beginners*. Retrieved September 28, 2021, from National Centre for  
17 Research Methods: <https://eprints.ncrm.ac.uk/id/eprint/245>
- 18  
19  
20 Pawson, R. (2013). *The Science of Evaluation: A Realist Manifesto*. London: Sage.
- 21  
22 Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.
- 23  
24 Petrosino, A., Boruch, ., R., Soydan, H., Duggan, L., & Sanchez-Meca, J. (2001). Meeting the  
25 Challenges of Evidence-Based Policy: The Campbell Collaboration. *BMJ*, 578 (1), 14-34.  
26 doi:10.1177/000271620157800102
- 27  
28  
29 Petrosino, A., Turpin-Petrosino, C., & Guckenburg, S. (2010). Formal system processing of juveniles:  
30 Effects on delinquency. *Campbell Systematic Reviews*.
- 31  
32  
33 Petrucci, C. J. (2021). If we measure it, they will come: A realist evaluation approach in a therapeutic  
34 jurisprudence context. *Family Court Review*, 59(3), 521–533. doi:10.1111/fcre.12592
- 35  
36  
37 Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. (1995). The well-built clinical  
38 question: A key to evidence-based decisions. *ACP Journal Club*, 123(3), A12-13.  
39 doi:10.7326/ACPJC-1995-123-3-A12
- 40  
41  
42 Sherman, L., Gottfredson, D., MacKenzie, D., Eck, J., Reuters, P., & Bushway, S. (1998). *Preventing  
43 Crime: What Works, What Doesn't, What's Promising*. Washington DC: US Department of  
44 Justice.
- 45  
46  
47 Sidebottom, A., Tompson, L., Thornton, A., Bullock, K., Tilley, N., Bowers, K., & Johnson, S. (2018).  
48 Gating Alleys to Reduce Crime: A Meta-Analysis and Realist Synthesis. *Justice Quarterly*, 55-  
49 86.
- 50  
51  
52 Sneddon, H., Gojkovic Grimshaw, D., Livingstone, N., & Macdonald, G. (2020). Cognitive-  
53 behavioural therapy (CBT) interventions for young people aged 10 to 18 with harmful sexual  
54 behaviour. *Cochrane Database of Systematic Reviews*.
- 55  
56  
57 Stephenson, M., & Allen, R. (2013). *Youth Justice: Challenges to practice*. Norwich: Unitas.
- 58  
59  
60 Sutton, C., Monaghan, M., Case, S., Greenhalgh, J., & Wright, J. (2022). Contextualising Youth Justice  
Interventions: Making the Case for Realist Synthesis. *Sustainability: Special Edition "Youth  
Justice Research"*, 14(2). doi:10.3390/su14020854

- 1  
2  
3 The Cochrane Collaboration. (2011). *Cochrane Handbook for Systematic Reviews of Interventions*,  
4 *Version 5.1.0*. Retrieved from Cochrane Handbook : <http://handbook-5-1.cochrane.org/>  
5  
6  
7 Thornton, A., Sidebottom, A., Belur, J., Tompson, L., & Bowers, K. (2019). On the development and  
8 application of EMMIE: insights from the What Works Centre for Crime Reduction. *29*(3).  
9  
10 Tompson, L., Belur, J., Thornton, A., Bowers, K. J., Johnson, S. D., Sidebottom, A., . . . Laycock, G.  
11 (2020). How Strong is the Evidence-Base for Crime Reduction Professionals? *Justice*  
12 *Evaluation Journal*, *4*(1).  
13  
14  
15 Turnbull, G. and Spence, J. (2011). What's at Risk? The proliferation of risk across child and youth  
16 policy in England. *Journal of Youth Studies*, *14*(8), 939-959.  
17 doi:10.1080/13676261.2011.616489  
18  
19  
20 Weisburd, D., Farrington, D., & Gill, C. (2017). What works in crime prevention and rehabilitation:  
21 An assessment of systematic reviews. *Criminology & Public Policy*, *16*(2), 415-49.  
22  
23  
24 Wilson, D. B., Brennan, I., & Olaghere, A. (2018). Police-initiated diversion for youth to prevent future  
25 delinquent behavior: A systematic review. *Campbell Systematic Reviews*.  
26  
27 Youth Endowment Fund. (2021, December 04). *Evidence and Gap Map of Interventions to Prevent*  
28 *Children Getting Involved in Violence*. Retrieved from Youth Endowment Fund:  
29 [https://youthendowmentfund.org.uk/wp-content/uploads/2021/05/YEF-map\\_10052021-1.html](https://youthendowmentfund.org.uk/wp-content/uploads/2021/05/YEF-map_10052021-1.html)  
30  
31 YJB (2021) *YJB Strategic Plan 2021-2024*. London: YJB.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Table 1: Summary of findings by EMMIE component**

<b>1. Cognitive Behavioural Therapy/CBT (Sneddon et al, 2020)</b>	
<b>Effectiveness:</b>	no info elicited from RCTs, poor comparison of intervention effects; bias via attrition; selective reporting; lacked blinding; incomplete outcome data, experimental research lacks rigour.
<b>Mechanisms:</b>	no reference to theory; simple black box explanations
<b>Moderators:</b>	no reference to contexts/moderators activating mediators/mechanisms
<b>Implementation:</b>	no discussion
<b>Economics:</b>	no quantification/estimation
<b>2. Juvenile Drug Courts (Mitchell et al, 2012)</b>	
<b>Effectiveness:</b>	small effect on reducing (re)offending; more rigorous experimental evaluations showed fewer effects; reduced adult reoffending more than juveniles'; evidence quality and internal validity concerns for quasi-experimental designs (e.g. selection bias)
<b>Mechanisms:</b>	no significant differences in reducing reoffending; court demands (e.g. frequent testing, hearings, monitoring) caused unintended consequences for children
<b>Moderators:</b>	adult-child differences linked to contextual differences in offence severity; higher recidivism reductions with non-violent child offenders
<b>Implementation:</b>	systematic documentation, insufficient detail; courts dealing with violent offenders less effective
<b>Economy:</b>	no quantification/estimation
<b>3. Formal system processing of juveniles (Petrosino et al, 2010)</b>	
<b>Effectiveness:</b>	processing (prevalence, incidence, severity, self-report outcomes) increases (re)offending likelihood; evidence quality concerns: inconsistent control conditions across studies
<b>Mechanisms:</b>	labelling as criminogenic; diversion effective when children view as onerous/deterrent or effective link to rehabilitative services
<b>Moderators:</b>	theory-based pre-specification of expected moderators/mediators relevant to activation of mechanisms; system processing had more positive effect for first-time offenders
<b>Implementation:</b>	systematic documentation, insufficient detail (e.g. diversion services ignored)
<b>Economy:</b>	no quantification/estimation
<b>4. Restorative justice conferencing (Livingstone et al, 2013)</b>	
<b>Effectiveness:</b>	no difference compared to court; poor evidence quality (bias, small numbers, lacked random allocation)
<b>Mechanisms:</b>	detailed theory articulation; reintegration/forgiveness (remorse, recognising wrongdoing, satisfaction)
<b>Moderators:</b>	contextual variations (e.g. 'recognition of wrongdoing')
<b>Implementation:</b>	no account; selection criteria may actively exclude implementation differences
<b>Economy:</b>	no quantification/estimation
<b>5. Police-initiated diversion (Wilson et al, 2018)</b>	
<b>Effectiveness:</b>	reduced offending (44% after year), but no other positive effects; no meaningful effect differences between diversion types; only high-quality experimental designs included
<b>Mechanisms:</b>	program theory - minor delinquency as normative activity in adolescence - exposure to deviant peers and labelling through system processing trigger offending
<b>Moderators:</b>	no differences by diversion type (e.g. only, referral to services) or offence type
<b>Implementation:</b>	systematic documentation, but poor description of delivery, lack of info on recidivism, client compliance, children's perceptions; biased police decision-making
<b>Economy:</b>	no quantification/estimation
<b>6. Multisystemic therapy for social, emotional, behavioral problems (Littell et al, 2005)</b>	
<b>Effectiveness:</b>	no significant effect on offending or other outcomes (behavioural, psychosocial, living arrangements, family) compared with other services/control; data of variable
<b>Mechanisms:</b>	no reference to theory; simple black box explanations
<b>Moderators:</b>	ad hoc description and too few studies in review for moderator analysis
<b>Implementation:</b>	ad hoc discussion; limited outcomes due to short-termism of interventions
<b>Economy:</b>	no quantification/estimation

Safer Communities - Decision on Manuscript ID SC-04-2022-0014

We would like to thank both referees for their detailed and diligent comments, which have been extremely helpful in strengthening this paper. We have sought to respond to the feedback and we explain the changes made below:

## Referee: 1 - Recommendation: Major Revision

**Comments:** This is a good paper, but the argument needs to be much clearer from the top. I think the inclusion of an introduction that introduces key terms, defines them and clearly outlines what your argument is would facilitate this. At present, the paper is a little too technical and needs properly introducing. Don't forget - you have submitted to a journal that is for academics, policy makers and practitioners. You have some important and valid points here, but if you do not properly explain and introduce them, there is a risk that the reader will get lost and not give the paper the attention it deserves.

**Response:** A detailed introduction section has been added, to respond to these comments, which outlines and defines terms and summarises the main arguments contained in the piece to make this clearer and more accessible to a range of stakeholders (also guided by point 6 below).

**1. Originality: Does the paper contain new and significant information adequate to justify publication?:** Yes - and suggests alternative frameworks by which to assess the effectiveness of youth justice interventions.

**Response:** We are pleased to hear this, thank you for noting.

**2. Relationship to Literature: Does the paper demonstrate an adequate understanding of the relevant literature in the field and cite an appropriate range of literature sources? Is any significant work ignored?:** Yes - a good knowledge and awareness of the literature is discussed.

**Response:** Thank you for acknowledging this.

**3. Methodology: Is the paper's argument built on an appropriate base of theory, concepts, or other ideas? Has the research or equivalent intellectual work on which the paper is based been well designed? Are the methods employed appropriate?:** Above, it states that this is an opinion paper. I am not sure I agree with this as the paper appears to be based on a systematic review. I would therefore suggest that this is a research paper. Some of the limitations are discussed. I would have liked to see more discussion of the limitations of applying adult based methods to evaluate interventions programmes based on adulterised understandings of youth crime.

**Response:** We have now embedded a critique of the responsabilising and adulterising nature of risk-led interventions and the 'What Works' evaluation framework that enables them, offering a number of explicit contrasts with the emerging 'Child First', child-centric, contextualised agenda for youth justice that recommends more contextualised ways of understanding and responding to children who offend through interventions and their evaluation.

**4. Results: Are results presented clearly and analysed appropriately? Do the conclusions adequately tie together the other elements of the paper?:** I think some more clarity is needed here. The title of the paper suggests that you are advocating for EMMIE over What Works. However, by the end of the paper you appear to be advocating for a realist synthesis approach to evaluation. This is not clear from the title, abstract or introduction. As stated below, I think more work needs to be done on properly introducing the paper to a reader who does not have the requisite expertise and knowledge of different evaluation methods. I think the argument that you are making also needs to be clearer at the outset. A clear introduction would facilitate this.

Safer Communities - Decision on Manuscript ID SC-04-2022-0014

**Response:** We have now amended the title, added detail to the introduction section, and amended the wording in the section 'Contextualising evaluations of youth justice intervention: Realist methods' to directly address these points.

**5. Practicality and/or Research implications: Does the paper identify clearly any implications for practice and/or further research? Are these implications consistent with the findings and conclusions of the paper?:** Yes - they are clear in the conclusion.

**Response:** Thank you for noting this. We are pleased to hear that these are clear.

**6. Quality of Communication: Does the paper clearly express its case, measured against the technical language of the field and the expected knowledge of the journal's readership? Has attention been paid to the clarity of expression and readability, such as sentence structure, jargon use, acronyms, etc.:** The paper launches into its arguments without firstly setting the context, and fully explaining what EMMIE, What Works and Realist Synthesis evaluation is and how it emerged. This needs to be at the top of the paper before the key arguments are discussed. It would also be helpful to define what you mean by 'effectiveness'. Are you referring to how it is defined in the legislation - prevention and reduction of reoffending (The Crime and Disorder Act, 1998), or something else? I feel that you need to firmly set the foundations and the context of the paper for the reader and this is missing at present.

**Response:** As reflected above, we have now added relevant detail to the introduction section to address each of these points.

## Referee: 2 Recommendation: Minor Revision

**Comments:** This is an interesting consideration of the limits of what works and EMMIE evaluations. The methodology is strong and it is intelligently analysed and considered. The literature used is appropriate for what it claims to do. However, there are a few areas where it needs to be strengthened:

1. it fails to address/recognise the move from a risk based system to a child first system of youth justice and what this means for the types of evaluation under review.
2. It discusses 'contextual' issues as being essential for evaluation and understanding and suggest a realist synthesis could capture change more effectively and allow us to understand what works for whom in what situations and why. However, there is little true consideration of realist synthesis and its strengths and weaknesses. In particular, is it valid in a child first system? Furthermore, are the questions now the same, does a child first system throw up different considerations?

**Response:** Both the introduction and the discussion sections have now been amended to include discussion of the move from a risk-based, adult-centric to a more contextualised and child-centric child first system and the consequent need for more contextualised interventions and evaluation methodologies. The discussion now has a more detailed section on the context-sensitive nature of the child first principle/tenets and how this has the potential to be more compatible and synergistic with a contextualised realist synthesis evaluation approach.

Safer Communities - Decision on Manuscript ID SC-04-2022-0014

1 **2. Relationship to Literature: Does the paper demonstrate an adequate understanding of the relevant literature in**  
2 **the field and cite an appropriate range of literature sources? Is any significant work ignored?:** The literature  
3 concerning evaluation and what works is all intelligently considered. However, the YJB and YOTs have rightly turned  
4 away from a wholly risk model towards a child first model in which risk is a reduced issue. Whilst 'what works' and  
5 EMMIE evaluations have their place in a risk model environment it is unclear how it has its place in a child first  
6 environment. The authors need to deal with the move away from risk and take that on board in assessing whether  
7 'what works' and EMMIE are still appropriate in a child first environment.  
8  
9

10  
11  
12 **Response:** Please see above. In the revised discussion, the new extended, child first-related consideration and  
13 justification of realist synthesis evaluation builds on the foregoing evaluation of the inapplicability of what works and  
14 the (more limited) applicability of EMMIE.  
15  
16

17  
18 **3. Methodology: Is the paper's argument built on an appropriate base of theory, concepts, or other ideas? Has**  
19 **the research or equivalent intellectual work on which the paper is based been well designed? Are the methods**  
20 **employed appropriate?:** The methodology is well designed for a risk based model of youth justice but we are no  
21 longer in that environment and the work needs to take account of child first.  
22  
23

24  
25 **Response:** Please see above  
26  
27

28  
29 **4. Results: Are results presented clearly and analysed appropriately? Do the conclusions adequately tie together**  
30 **the other elements of the paper?:** Yes, within a risk based paradigm the results are intelligently and fully analysed.  
31 The article claims that realist synthesis can better evaluate why, how, where and for whom an intervention might  
32 work/fail to work. However, I would have liked to see more explanation of the realist synthesis before coming to this  
33 conclusion and particularly, again looking at what might work in a child first system.  
34  
35

36  
37 **Response:** The emergence and nature of realist synthesis are now discussed in more detail in the introduction, whilst  
38 its applicability (when compared to EMMIE) and its contextualised benefits in a child first environment now asserted  
39 in the revised discussion section.  
40  
41

42  
43  
44 **5. Practicality and/or Research implications: Does the paper identify clearly any implications for practice and/or**  
45 **further research? Are these implications consistent with the findings and conclusions of the paper?:** It does, if we  
46 were still in a risk based system it would be fine. What it fails to do is to consider is the effects of a move from a risk  
47 based system to a child first system for the types of evaluations being considered.  
48  
49

50  
51  
52 **Response:** Please see above  
53  
54  
55  
56  
57  
58  
59  
60