

AN INVITATION TO TENDER FOR A SYSTEMATIC REVIEW OF THE OUTCOME MEASURES ASSOCIATED WITH YOUTH CRIME AND VIOLENCE

August 2020

The Youth Endowment Fund (YEF) is commissioning a systematic review of how best to measure the intermediate and ultimate outcomes for youth crime and violence.

Background

Early intervention is essential if we are to tackle youth crime and violence over the long term. The Youth Endowment Fund is a ten-year investment of £200m to fund, support and evaluate programmes and local partnerships working with 10–14 year olds in England and Wales to find out what works. The YEF does this by:

- Supporting and evaluating the most promising interventions aiming to prevent 10–14 year olds in England and Wales from being drawn in to crime and violence;
- Building the UK evidence base of what works by generating and promoting knowledge and practice to transform how we tackle youth offending;
- Developing a 'place based' approach to working with local community partnerships in selected areas of high need; and
- Building partnerships with other funders in order to share best practice and learning while directing more resources to tackling serious violence and building a legacy beyond the ten-year life of the fund.

In order to build the evidence base for what works in tackling youth offending, the YEF will commission evaluations of programmes with the aim of reducing youth crime and violence, using reliable and valid outcome measures.

To date the YEF has commissioned 24 evaluations worth approximately £3m and the Covid-19 Learning Project. Outcome measures have been selected based on the intervention's theory of change and ours and evaluators' knowledge of high-quality and commonly used measures such as the Strengths and Difficulties Questionnaire and Self-Report Delinquency Scale.¹² Going forward the YEF would like to be more systematic in identifying and selecting the most reliable and valid outcome measures for its youth justice interventions.

The need for this review

Ultimately the YEF wants to know whether the interventions it evaluates reduce crime and antisocial behaviour and intends to assess this via long-term follow up of official records of offending behaviour on the Police National Computer. However, given the YEF's focus is on young people aged 10 to 14 – the majority of whom will not yet be involved in serious or violent crime – and that many of the interventions will have a primary focus on addressing upstream risk and protective factors associated with later offending – such as school exclusions, externalising behaviours or family function – it is important to identify:

- 1) Reliable and valid measures of the causal mechanisms and intermediate outcomes (or 'risk and protective factors') that might lead to a reduction in later violence and crime; and
- 2) Reliable and valid self-report measures of violence and crime.

In 2015, the Early Intervention Foundation (EIF) commissioned a review of the risk and protective factors for young people aged seven to 25 associated with involvement in youth violence and gangs, using correlational evidence from longitudinal studies.³ The individual factors most strongly associated with youth violence included impulsivity, low self-esteem and aggressive behaviour.

¹ Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581–586.

² <https://www.edinstudy.law.ed.ac.uk/>

³ <https://www.eif.org.uk/files/pdf/preventing-gang-and-youth-violence-risk-protective-factors.pdf>

However, the review did not identify which factors are causally associated with crime and violence. There are also many different measures available of these factors, of varying quality. For example, in 2005 the National Centre for Injury Prevention and Control of the Centres for Disease Control and Prevention in the U.S published a compendium of over 170 measures of violence-related attitudes, behaviours and influences amongst youths.⁴

This project will build on these pieces of work, and the work being conducted by the YEF in Annex 1, by reviewing the measures that causally predict youth violence and crime in 10 to 14 year olds, in the UK, using research from experimental or quasi-experimental studies. It will also identify the best available measures of self-report violence and crime. Information on quality and psychometric properties (for example, using EIF's approach in Annex 2) will be extracted to inform a database of YEF measures, with the aim of supporting YEF evaluators to choose the highest quality, most comparable measure for each study and to avoid variation in decision making across projects.

Aims of the review

The YEF anticipates that there are four strands to this review:

- 1) Map theoretical constructs (including individual-level risk and protective factors) associated with crime and violence amongst 10 to 14 year olds.
This work will initially be done in-house by YEF.
- 2) Systematically review the available measurement tools of those constructs.
- 3) Review the extent to which those measurement tools are malleable and causally predictive of crime and violence.
- 4) Review the quality of these measures based on an agreed set of criteria, to inform a database of measures.

⁴ https://www.cdc.gov/violenceprevention/pdf/yv_compendium.pdf

Annex 1 of this ITT summarises the work to date that has been completed with respect to strand one. The YEF is anticipating that the review will build on the work in Annex 1 and also complete strands two to four. The YEF is open to discussing the optimal order in which the strands should be completed and the scope of the review.

Approach

1) Mapping theoretical constructs (including individual-level risk and protective factors) associated with crime and violence.

Annex 1 summarises the YEF's work to date. Given the potential breadth of scope of the review, the YEF will want the successful bidder to prioritise the identification of measures of individual-level factors that are most closely aligned with YEF's ultimate outcome of youth crime and violence (e.g. aggression, rule-breaking) and that could be primary outcomes in YEF evaluations, as opposed to more upstream individual factors (e.g. self-efficacy or internalising problems) that could be secondary outcomes in YEF evaluations. The YEF will finalise the review scope in discussion with the appointed reviewer.

Once the scope has been agreed, the reviewer may need to conduct a concept mapping exercise using theoretic and empirical literature to identify the various theoretical constructs and terminology used to describe them, to inform the search terms for the subsequent review. However, the YEF expects the majority of the reviewer's time to be spent on aims 2-4.

2) Systematically reviewing the available measures of those constructs.

Given the importance of this work, the YEF's preferred approach is to commission a full systematic review (see Box 1).

There are likely to be several stages to the filtering process for measures. The first stage may, for example, involve identifying all measures of the risk and protective factors in scope that are written in English, report sufficient information on

psychometric properties, take less than an hour to administer, and are obtainable (e.g. if not free, can be purchased).

Box 1: The case for a systematic review

The strengths of systematic reviews are well-documented: they provide the best available evidence, on a research question, which is identified, selected, appraised and synthesised in a transparent or unbiased way.* Full systematic reviews produce higher quality results than rapid evidence assessments and tend to include the following features:**

- Clear research questions;
- Transparent inclusion criteria;
- A published review protocol, subjected to peer review;
- Searches that attempt to locate all relevant published and unpublished studies to limit impact of publication biases and use methods to reduce bias in study selection (e.g. using double screening);
- Data extraction that is determined by the research questions, and may include methods to reduce bias in data extraction (e.g. double independent extraction);
- Critical appraisal that examines in a systematic manner the methods used and investigates potential biases in those methods;
- Transparent methods to synthesise and reach conclusions.

* The Methods Group of the Campbell Collaboration, 2017. Methodological expectations of Campbell Collaboration intervention reviews: Conduct standards. Campbell Policies and Guidelines Series No. 3. DOI: 10.4073/cpg.2016.3.

** Petticrew, M. (2001). Systematic reviews from astronomy to zoology: myths and misconceptions. British Medical Journal, 322(January), 98-101.

3) Review the extent to which these measures are malleable and causally predictive of crime and violence.

Once a long-list of measures has been identified. The YEF anticipates that the reviewer will need to conduct a second systematic review, to identify evidence for

the causal relationship between these measures and crime and violence, using evidence from experimental and quasi-experimental studies. This will enable the reviewer to identify not just whether these measures are correlated, but also those that are most malleable and causally predictive of, crime and violence.

Reviewers may want to consider using a similar approach to that used by the Institute of Education in its review of the impact of ‘non-cognitive skills’ on attainment and employment outcomes for young people, for the Education Endowment Foundation.⁵ In this review each factor was assessed for malleability, effect on other outcomes, and strength of the evidence.

Where insufficient evidence is available, the review should identify what future research would be most valuable. There may be an argument for completing this stage before stage two and YEF is open to discussing this with the appointed reviewer.

4) Review the quality of these measures based on an agreed set of criteria, to inform a database.

The final part of this project will involve systematically extracting data on each measure in the final list, based on an agreed set of coding criteria. For this strand of the work, reviewers will be expected to build on the EIF measurement assessment criteria (see Annex 2) and the approach used to develop the EIF’s SPECTRUM database.^{6 7}

The criteria will include (but is not restricted to) the following information:

- The specific construct the instrument measures, its conceptual relationship with the causal mechanisms or intermediate outcomes identified in the first part of the review, and its subscales;

⁵ https://educationendowmentfoundation.org.uk/public/files/Presentations/Publications/Non-cognitive_skills_literature_review_1.pdf

⁶ <https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluating-projects/measuring-essential-skills/spectrum-database/>

⁷ https://educationendowmentfoundation.org.uk/public/files/Evaluation/SPECTRUM/V6_Guidance_document.pdf

- History of use and availability;
- Psychometric properties including reliability, validity, and responsiveness to change (including sample sizes, country and year of tests);
- Details of the sample(s) that the measure has been tested with;
- Strength of the evidence that it causally predicts crime and violence, if not already covered under predictive validity;
- Implementation properties and practical considerations, such as requirements for administration, instrument format, age range, time and cost;
- Readiness to use in the UK.

It is anticipated that these criteria would be input into a spreadsheet of measures, which could subsequently be converted into a front-facing static or searchable online database by the YEF. The database will need to have a way of summarising overall quality and suitability of each measure, for example using a rating scale, and the YEF may want identify a 'preferred measure' for each domain.

Self-report measures of violence and crime

The YEF is keen to include measurement of violence and crime in the review. Violence and crime as measured by administrative records in the Police National Computer (e.g. arrests and convictions), represents only a fraction of violence and crime that occurs. The YEF is interested in exploring best practice and identifying the best available instruments for measuring self or other-reported violence and crime, and minimising bias, to support its evaluators to deliver informative and comparable impact evaluations.

To date the YEF has used and produced guidance on the Self-Report Delinquency Scale that was developed for Sweep 3 of the Edinburgh Study of Youth Transitions

and Crime.⁸ This results in two measures; one of volume and one of variety of delinquency, which poses challenges for interpreting intervention impact. For example, is a reduction in severity but an increase in volume of delinquent behaviour a positive, negative or null result? We are interested in bidders' views on how best to approach this part of the review.

Outputs and timeframe

The YEF will work to finalise the specification with the successful team and the resulting protocol will be peer reviewed before being published on the YEF's website.

The YEF anticipates that the final outputs will include:

1. A conceptual piece on the constructs included in the review, and rationale, possibly co-authored between YEF and the reviewer;
2. An Excel database of measures summarizing all the extracted criteria;
3. A brief guidance document on best practice in measure choice, implementation, analysis and interpretation, for evaluators and practitioners, to accompany the database;
4. A technical appendix, fully documenting the rationale, methods and results of the reviews.

The YEF is also interested in discussing whether there may be any opportunities for interim publications, other than those suggested above. The YEF would like applicants to consider the minimum time it would take to produce a high quality set of outputs. We expect the review of priority constructs to be completed within a year of commissioning.

⁸ <https://www.edinstudy.law.ed.ac.uk/>

How to apply

The process will involve two stages:

1) Expression of interest

Applicants should submit a short expression of interest to the YEF of no more than 1000 words. The expression of interest should include:

- Relevant experience of the project team and credentials for undertaking the work, including:
 - Knowledge and experience of delivering and/or reviewing youth justice impact evaluations;
 - Knowledge and experience of outcome measurement in this field;
 - Experience in conducting systematic reviews;
- A brief summary of the overall approach that would be taken and key considerations for the design of the review.

The YEF will review the EOIs based on the criteria provided in Annex 3. **Please send your expression of interest to amy.wells@youthendowmentfund.org.uk by 5pm 18th September 2020.**

2) Proposal stage

The YEF expects to request a maximum of three teams to prepare proposals of no more than 4000 words, excluding references and are aiming to notify bidders on the 25th of September. In the proposal, bidders will be expected to include the following:

- Objectives of the review, including research questions and rationale for why these are important;
- Your proposed approach to conducting the review, including a detailed description of how the theoretical constructs and their instruments will be identified, selected and appraised (including clear inclusion criteria);

- For the measures database, please specify the information that can be made available using your approach, and how this will help evaluators and practitioners select appropriate measures;
- The relevant experience of the team. Specifically, your credentials for undertaking the review, your knowledge in the subject matter and expertise in the proposed methodology;
- A detailed budget with justification for time spent on the research.

Given the level of ambiguity around the scope (i.e. number of constructs to be included) the YEF would welcome bidders providing budgets based on the assumption of 10–15 constructs, 15–20 constructs and 20–30 constructs being included in the review.

Bidders will be given three weeks to prepare proposals and the deadline is likely to be 5pm on the 16th of October. Bidders will be able to submit questions by the 2nd of October. All proposals will be reviewed by the YEF, using the criteria provided in Annex 3. The YEF is likely to want to have follow up interviews with the preferred bidders during the week of the 19th of October, before making a final decision. The YEF is happy to consider bids from consortiums, as long as consideration is given to communication and how the teams will work well together. The YEF would expect to contract with one lead member of the consortium.

Annex 1: Mapping constructs and their associations with crime and violence

The YEF Evidence and Evaluation team are carrying out an exercise ‘in house’ to review the constructs associated with risk and protective factors for crime and violence, focusing first on individual level constructs. It is expected that the commissioned review will build on this work. The YEF will share the work it has done to date with the preferred bidder.

The aim of this work is to have a better understanding of the relationship between key constructs and later offending, violence and antisocial behaviour.

This work is not intended to be systematic but focuses on key literature. For each construct we are identifying and extracting information from:

1. One longitudinal study or review of longitudinal studies, to show whether or not the construct predicts crime and violence and the magnitude of the association;
2. One intervention study (e.g., from an RCT or QED) or review of intervention studies to show whether an intervention that changes the construct at time 1 also reduces offending at time 2. Moreover, does the change in construct at time 1 predict the reduction in offending in time 2?

See Table 1 for an example. For both longitudinal and intervention evidence, systematic reviews and meta-analyses are prioritised. Longitudinal literature was identified predominantly by using the YEF’s Initial Evidence Review, the Campbell Collaboration⁹ and Cochrane Library Website¹⁰ and intervention literature was identified, in the first instance, using the EIF guidebook¹¹, Crime Solutions¹² and Blueprints¹³.

⁹ <https://campbellcollaboration.org/>

¹⁰ <https://www.cochrane.org/evidence>

¹¹ <https://guidebook.eif.org.uk/>

¹² <https://crimesolutions.ojp.gov/>

¹³ <https://www.blueprintsprograms.org/>

1. Table 1. An example of the mapping associations exercise

Construct: Bullying perpetration	
Longitudinal evidence	Intervention evidence
<ul style="list-style-type: none"> • Systematic review with the aim of investigating the relationship between bullying perpetration and later offending • Self-report and official records of offending and delinquency • The probability of offending up to 11 years later was much higher for school bullies than for non-bullies (OR = 2.50] • Bullying perpetration was a significant risk factor for later offending even after controlling for major childhood risk factors [OR = 1.82] 	<ul style="list-style-type: none"> • A QED (extended age cohort design; EIF guidebook level 3) was used to evaluate a school-based, anti-bullying intervention in 42 schools over a two year period • N = 2500 children • Self-report antisocial behaviour • The intervention group showed a significant reduction in bullying perpetration <i>and</i> in antisocial behaviour (including vandalism and fighting)

Indicative list of constructs

From the work outlined above (mapping constructs and their associations with crime and violence) we have compiled an **indicative** list of constructs that could be included in the systematic review of measures (please see below). Please note that the successful bidder will need to work with YEF to finalise the scope of the review and to determine the final list of constructs to be included. For example, priority should be given to those constructs that could be primary outcomes in YEF evaluations.

Indicative list of priority constructs, based on work to date:

- Delinquency
- Antisocial behaviour
- Weapon carrying
- Externalising problems
- Anger
- Aggression
- Bullying perpetration
- Gang involvement
- High daring
- Impulsivity
- Alcohol use/misuse
- Substance use/misuse
- Callous-unemotional traits

Other constructs that could be considered:

- Delinquent peers
- Conflict resolution style
- Internalising problems
- Attitudes towards
 - o Offending
 - o Alcohol and drugs
 - o School
- Future aspirations
- Prosocial attitudes
- School attendance/truancy
- School exclusion
- Low attainment
- Low engagement
- Resilience
- Hope
- Attachment
- Mental health and wellbeing

- Empathy
- Self-esteem

Annex 2: EIF measurement assessment criteria

In 2019 EIF conducted a rapid evidence assessment of measures used to measure parent conflict and its impact on child outcomes¹⁴. In doing so EIF developed measurement assessment criteria that the YEF would expect the reviewer to build upon when developing the criteria for inclusion in the YEF measures database. Measures were assessed against psychometric and implementation measures. Please see the table below for details of the measurement assessment criteria. Further information can be found in [EIF's report](#) Appendix C (p.113).

Features	Description	Criteria
Psychometric features		
Internal consistency	The degree to which items designed to measure the same outcome relate to one another.	Cronbach's alpha coefficients ≥ 0.60 .
Test-retest reliability	The extent to which the outcomes of an assessment are stable over time. Assuming no important intervening events, a person's score on a measure taken multiple times should be correlated.	Test-retest scores (for example ICC) ≥ 0.70 for short periods of time (1–4 weeks).

¹⁴ Pote et al (2020). Measuring parental conflict and its impact on child outcomes. Guidance on selecting and using valid, reliable and practical measures to evaluate interventions. <https://www.eif.org.uk/resource/measuring-parental-conflict-and-its-impact-on-child-outcomes>

<p>Validity</p> <p>Criterion validity</p> <p>AND/OR</p> <p>Construct validity</p> <p>AND/OR</p> <p>Concurrent validity</p>	<p>Criterion validity: The extent to which scores on a particular measure are related to a 'gold standard'.</p> <p>Construct validity: The extent to which scores on a particular measure relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the measured concepts.</p> <p>Concurrent validity: The extent to which scores on a new measure are related to scores from a criterion measure administered at the same time.</p>	<p>Criterion validity is reported, with a convincing argument that the gold standard is 'gold' and the correlation with that gold standard (typically AUC) is ≥ 0.70</p> <p>AND/OR</p> <p>Pearson correlation tested against similar measures $\geq 0.70^*$</p> <p>AND/OR</p> <p>Confirmatory Factor Analysis shows strong results (CFI ≥ 0.95 and RMSEA < 0.05).</p>
<p>Sensitivity to change</p>	<p>The ability of a measure to detect important changes over time (post-intervention), even if these changes are small.</p>	<p>The measure has been used and found to identify change over time in two impact evaluations (for example pre-posts/QEDs/RCTs) of relevant interventions conducted over a short period of time.</p>
<p>Implementation features</p>		

Brevity	The time taken to complete a measure and/or the number of items in a measure.	The measure takes ≤ 15 minutes to complete AND/OR It comprises of ≤ 30 items.
Availability	The extent to which a measure is freely available.	The measure is free to use.
Ease of scoring	The extent to which a measure is easy to score and interpret.	The measure has simple scoring instructions AND It does not need to be scored/ interpreted by someone with specific training or qualifications.
Used in the UK	The extent to which the measure has been used in the UK setting.	The measure has been used in at least two UK studies published in peer-reviewed journals by different research teams.

Note: *If < 0.70 , case-by-case decisions are made.

Where this information is available, the reviewer may also want to look at:

- Whether the measures are causally predictive of violence and offending.
- Freedom from biases (e.g. by gender, social class or ethnicity).
- Range, including freedom from ceiling and floor effects.
- Robustness (i.e. is the test 'objective' in the sense that it cannot be influenced by the expectations or desires of the assessor).
- Administration requirements (e.g. does it need to be invigilated or administered by a qualified person?).

Annex 3: YEF scoring criteria

EOI scoring criteria

1. Relevant experience of core project team and understanding of topic area (50%)

- a. The extent to which the applicant demonstrates an understanding of youth justice impact evaluation and outcome measurement.
- b. The extent to which the proposed team demonstrates a track record of delivering systematic reviews.

2. The suitability of the proposal overall approach and key considerations identified for the review (50%)

Proposal scoring criteria

1. Relevant experience of core project team and understanding of topic area (30%)

- a. The extent to which the applicant demonstrates an understanding of youth justice impact evaluation and outcome measurement.
- b. The extent to which the proposed team demonstrates a track record of delivering systematic reviews.

2. Methodology and approach (40%)

- a. The suitability of the proposed scope of the review and understanding of the research aims and objectives.
- b. The suitability of the proposed approach to identify, select and appraise the available measures, to deliver against the research aims and objectives.
- c. The suitability of the proposed approach to identify, select, and appraise the available evidence of a causal link or association between measures and youth crime violence, to deliver against the research aims and objectives
- d. The suitability of the information that will be made available in the measures database and approach proposed to judge measurement quality.

3. Value for Money (30%)

a. Detailed cost of your proposal and how this demonstrates value for money

Scoring criteria	
0	Totally fails to meet the requirement - information not available
1	Meets some of the requirements with limited supporting information
2	Meets some of the requirements with reasonable explanation
3	Fully meets the requirements with detailed explanation and evidence
4	Exceeds the requirements with extensive explanation and evidence