Youth Justice Board Effective Practice Standards of Evidence: an update

Version 3

YJB Effective Practice Team
Youth Justice Board for England and Wales

Contents

Introduction

The 1998 Crime and Disorder Act set out key reforms for the youth justice system, forming the basis for a new youth justice framework. The Act also saw the establishment of the Youth Justice Board (YJB) which was given a range of statutory duties, one of these being 'to identify, to make known and to promote good practice[1] ...' The Standards of Evidence presented here play an important part in this role, by clearly categorising the strength of the evidence underpinning submitted examples of effective practice.

About this guide

This guide describes the updated YJB Standards of Evidence[2] that it uses to classify submissions made by practitioners to the Library of Effective Interventions. The Standards of Evidence presented here both update and replace the previous standards used by the YJB to assess submissions to its library.

The aim of the guide is to describe how the YJB, and its Effective Practice Classification Panel will assess the evidential strength of submissions of good practice. In doing so it also provides a guide to inform those practitioners or others wishing to submit examples.

We will use these Standards of Evidence to identity the highest quality approaches/interventions (in turn helping users of our Effective Practice materials understand which interventions they can have the most confidence in), while at the same time discovering those which would benefit from further evaluation.

What is the YJB Library of Effective Interventions?

The YJB's Library of Effective Interventions (or Effective Practice library, as it is often known) is an online collection of practice resources, materials and information, intended to:

- provide practitioners and commissioners in youth justice with easy access to examples of effective approaches/interventions from other services and providers; and
- allow those developing and using innovative practice to share what they have found to be effective in working with young people and delivering youth justice services[3].

We hope that the library is a helpful reference tool for practitioners and managers, who will be able to access a comprehensive collection of well-evidenced approaches/interventions in one place.

To facilitate this the library is incorporated within the wider online 'Youth Justice Resource Hub' set up by the YJB, which has brought together a number of its resources into one place. It is anticipated that fully integrating these resources will make it easier for users (particularly youth justice practitioners) to navigate and search through the materials the YJB holds.

---

[1] http://www.legislation.gov.uk/ukpga/1998/37/section/41/england/2002-12-02?view=plain
[2] These Standards of Evidence replace the former Youth Justice Board 'Practice Classification Framework'.
[3] https://yjresourcehub.uk/effective-practice.html

What are the Standards of Evidence and why have them?

Approaches/interventions catalogued on the library have been classified by the YJB's Effective Practice Classification Panel.

On the library are examples of approaches/interventions on similar themes. The classification of a particular approach/intervention will indicate how strong the supporting evidence for it is. The overall classification is based on an aggregate of the quality of evidence that has been provided to the panel by the submitter, in the following areas:

- The description of, and theory behind, the approach/intervention (what it is and why it worked, or should work).
- The approaches/intervention's effectiveness (how far it has worked).
- The context of the approach/intervention (where it worked and for whom).
- How much the approach/intervention cost and whether it has been cost-beneficial (whether or not it represented value for money).

Using clear standards of evidence allows others to assess how well-proven the examples of effective practice are – an important consideration in possible replication.

Care has been taken to ensure that the Standards of Evidence published in this guide correspond with those from existing organisations, such as those developed by the Social Research Unit at Dartington for Project Oracle, and those used by the Early Intervention Foundation - for example a submission classified at Level 1 by the YJB is similar to those classified at Level 1 by these organisations.

What to consider when reading examples on the library

When looking at the approaches/interventions listed on the library, it is important to remember the following:

1. The **context** of any approach/intervention is crucial as places differ greatly in terms of their social and physical environment. Only those examples listed at the highest levels on the library (levels 2 and 3) will include information about the context in which the approach/intervention operated. If you are considering implementing a particular approach/intervention in your area, you will need to think about whether and what adjustments might need to be made.

2. Appropriate **delivery** of any approach/intervention is critical. Central to this is whether it has been delivered as intended, with its key 'ingredients' adhered to (programme fidelity), but the skills of those delivering it are also important.

3. How **ready** the approach/intervention is to be implemented elsewhere or more widely. Different approaches/interventions will be at different stages of readiness. Some will be able to indicate how ready they are to implement in other areas (such as costs and resources required), while others will not have this information. The evidential level achieved by a submitted approach/intervention is not an indication of readiness for roll-out.

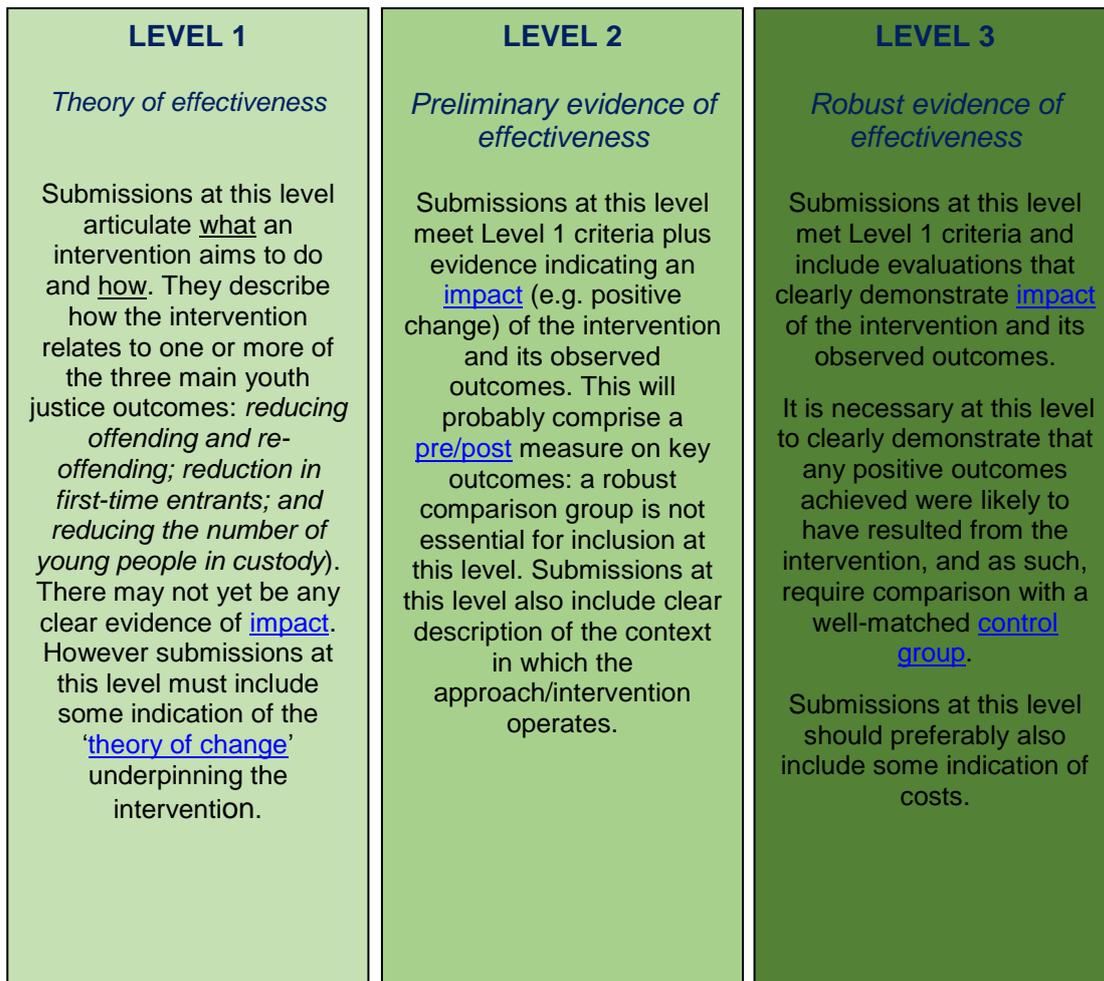## What to know if you are thinking about making a submission to the library

Before submitting an example for inclusion onto the library, you should make sure that you read these standards of evidence carefully. This will help you to guide your

submission. Once you have done this you will need to complete a submission form[4] and send it to the YJB Social Research and Effective Practice Team. On the form you will be asked to describe details of your approach/intervention, along with any indication of its effectiveness. Further details are also available from the above link to the form.

YJB Standards of Evidence - classification system

The Standards of Evidence used for the YJB Library of Effective Interventions are now comprised of Levels 1, 2 and 3 (the least to the most robust). These classifications replace the previous ones of 'Research-proven', 'Promising', and 'Emerging'.

This classification system is summarised in the diagram below, and described in more detail throughout this guide.

LEVEL 1	LEVEL 2	LEVEL 3
Theory of effectiveness	*Preliminary evidence of effectiveness*	*Robust evidence of effectiveness*
Submissions at this level articulate what an intervention aims to do and how. They describe how the intervention relates to one or more of the three main youth justice outcomes: *reducing offending and re-offending; reduction in first-time entrants; and reducing the number of young people in custody*). There may not yet be any clear evidence of impact. However submissions at this level must include some indication of the 'theory of change' underpinning the intervention.	Submissions at this level meet Level 1 criteria plus evidence indicating an impact (e.g. positive change) of the intervention and its observed outcomes. This will probably comprise a pre/post measure on key outcomes: a robust comparison group is not essential for inclusion at this level. Submissions at this level also include clear description of the context in which the approach/intervention operates.	Submissions at this level met Level 1 criteria and include evaluations that clearly demonstrate impact of the intervention and its observed outcomes. It is necessary at this level to clearly demonstrate that any positive outcomes achieved were likely to have resulted from the intervention, and as such, require comparison with a well-matched control group. Submissions at this level should preferably also include some indication of costs.

Youth Justice Board effective practice threshold

Not proven to be effective	*Proven to be ineffective or harmful*

---

[4] This link also provides further instructions about how to make a submission to the Library of Effective Approaches and Interventions.

Submissions that do not meet the YJB's effective practice threshold

The majority of the submissions that the YJB receives do not meet its effective practice threshold. In most cases this is because there is not enough evidence presented with them to measure effectiveness. However, in some cases this may be because the evidence presented with them suggests they are ineffective or harmful.

These submissions are described below.

Approaches/interventions will be regarded as '**Not proven to be effective**' if no evidence is provided, or an examination of evidence that does accompany them was not of sufficient quality to suggest effectiveness. Approaches/interventions at this level may still achieve Levels 1, 2 or 3 with more robust evaluation.

Other submissions however may have very robust evaluations attached to them, but the evidence they present suggests that they may be ineffectual. These submissions are regarded as '**Proven to be ineffective or harmful**'.

Approaches/interventions will be regarded as ineffective or harmful if they have been found, based on rigorous evaluation evidence (i.e. to a standard seen in Level 3) to have *consistently* neutral or negative effect on their observed outcomes.

The remainder of this guide is dedicated to those submissions meeting the YJB's effective practice threshold.

How to achieve the YJB Standards of Evidence

Level 1

This is the first level of evidence passing the YJB's effective practice threshold, Submissions at this level may be largely descriptive but must have a logical foundation or theory supporting the practice described. Without a clear idea of where we want to be, we do not know logically if the direction in which we are going is the right one. It is for this reason that this is the minimum threshold.

Submissions at this level describe what outcomes the approach/intervention intends to achieve and provides a clear and logical illustration of how that happens (or will happen). They clearly describe:

- what the situation was prior to the approach/intervention;
- what issue they are aiming to change;
- who will be eligible (target group)
- the approach/intervention to be implemented to change this; and
- the theory or process behind the choice of that approach/intervention.

A useful framework for this process is what is termed a theory of change. This is simply a method for describing what an approach aims to do and how it aims to do it. Being able to set out logically the theory for any approach/intervention in this way is the starting point of an evaluation and the collection of evidence to demonstrate its effectiveness.

If the intervention has already been implemented, then the submission will describe clearly what the intervention is, how it works, and with/for whom. A description of the approach/intervention should include the following information:
- content;
- duration;
- frequency;
- setting;
- mode of delivery (e.g. face to face or virtually); and
- format (e.g. group or individual).

What makes a good theory of change?

There are several ways to represent a theory of change and they can range from very basic to highly complex. However a good theory of change should be:

- Clear - It should describe what is done, how it is done and what is achieved, either in narrative or diagrammatic form.

- Demonstrable - It should be possible to test any assumptions that certain activities will lead to certain observed outcomes.

- Contextualised - It should contain a narrative that explains the background and/or circumstances to the approach/intervention[5].

If for example an approach/intervention uses a particular method that already has a well-established evidence base (e.g. Restorative Justice) and demonstrates how the principles of this method are applied in the approach/intervention with fidelity, this would satisfy the criteria for Level 1.

---

[5] Noble J and Hodgson L (2014) Theory of Change Guidance. Youth Justice Board for England and Wales.

The Youth Justice Board commissioned the company New Philanthropy Capital (NPC) 'to deliver training to practitioners in the youth justice sector on the value and practice of developing theories of change'. The guidance they produced is available here. As part of this process NPC also developed a series of webinars, one of which is concerned with theory of change and is available here.

<div style="border:1px solid #000; background:#9f8fb0; padding:1em;">

Evidence required against the four classification criteria used by the Classification Panel: Level 1

1. **Why it works -** A theory of change or logic model that clearly describes how the approach/intervention achieves its intended outcomes, and/or reference to an existing evidence base for the approach.

2. **Context** - A description of the context in which the approach/intervention operates, including information about:

   - Any referral or assessment processes.
   - Offender/participant, offence, gender, age etc.
   - Geographical location and its socio-demographic characteristics.
   - The agency or agencies involved in delivery.
   - The population served.
   - Evidence of a needs analysis that shows how the approach/intervention meets the assessed needs of the target group.

3. **Effectiveness** – not required

4. **Cost** - not required

</div>

For further (and more detailed) reading on developing a theory of change and on evaluation more generally, The Magenta Book is the recommended central government guidance that sets out best practice for its departments. It presents standards of good practice in conducting evaluations, and seeks to provide an understanding of the issues faced in this area.

Levels 2 and 3 - evidence of impact

The threshold between Level 1, and Levels 2 and 3 in the Standards of Evidence is where practice moves from theory-based design to a more developed evaluation stage. At Levels 2 and 3, submissions begin to contain evidence of impact of approaches/interventions. This is fundamental for inclusion at these levels.

Submissions at these levels satisfy all of the criteria needed for Level 1 but go on to describe outcomes, and in particular, whether the observed outcomes correspond to the expected outcomes.

Level 2

The requirement for classification at Level 2 is not that of the most robust or comprehensive evaluation design. However, some basic principles are important and do have to be met. Submissions at this level:

- Clearly describe the context in which the approach/intervention operates, but also provides insight into why it works (compared to submissions at Level 1). It will also include:
  - o a description of what it is and how delivered;
  - o more detailed characteristics of the target group; and
  - o any specific contextual factors or conditions required for it to be effective.

- Demonstrate the distance travelled since the start of the approach/intervention. For example a baseline measurement **must** have been taken at the start, against which progress has been calculated at the end (and ideally with a follow-up period extending beyond the period of the approach/intervention itself, to see if any benefits have been maintained over time). This style of evaluation is commonly known as a pre/post design. Evaluations at this level show empirical evidence that the approach/intervention has had an impact on an intended target group.

Examples of stronger submissions at Level 2 also include comparisons between those receiving the approach/intervention (the target group) and a group not receiving it, or some other relevant comparison. Also desirable at this level, although again not essential is information about how much the approach/intervention costs to deliver. These are however, the factors that need to be addressed by submissions in Level 3 of the Standards of Evidence.

> **Evidence required against the four classification criteria used by the Classification Panel: Level 2**
>
> 1. **What it is and why it works** - A theory of change/logic model that clearly describes how the approach/intervention achieves its intended outcomes, and/or reference to an existing evidence base for the approach.
>
> 2. **Context** - A thorough description of the context in which the approach/intervention operates, including all information about context from Level 1.
>
> 3. **Effectiveness** - A pre/post design that demonstrates the distance travelled for the target group against a baseline taken prior to the delivery of the approach/intervention (preferably with a post-intervention follow-up measure).
>
> 4. **Cost** – not required

Level 3

Submissions at Level 3 provide the most robust examples of evidence. They are distinguished from those at Level 2 by the ability to say with greater confidence that the approach/intervention is likely to have produced the observed outcomes.

To do this appropriate methods need to be used which incorporate comparisons between those who underwent the approach/intervention to those who didn't. These methods could include but are not limited to the utilisation of a well matched control group via:
- randomised allocation to the approach/intervention or a comparator group;
- regression discontinuity designs and comparison time series designs; or
- alternative designs (such as Propensity Score Matching)

The method must clearly and comprehensively demonstrate the impact of the approach/intervention being measured. This enables assessment of what would have happened to the outcomes of interest, if the approach/intervention had not been introduced. These evaluations are more comprehensive and robust, controlling for as many of the factors that can have an effect on the potential outcomes of an approach/intervention as possible.

We understand that robust outcome evaluations of this nature are likely to require the involvement of professional researchers (e.g. Universities or research organisations). However, the Ministry of Justice's Data Lab can also provide support to organisations to help them understand the outcomes of their work by providing 'an analysis which assesses the impact of that programme against various ... measures'[6].

While not necessary for inclusion into the library, commissioning an organisation to assess the impact of an approach/intervention can mean that findings are viewed as independent of the commissioning organisation. Maintaining this independence can sometimes be very important in considering the findings of an evaluation and assessing the relative success of an approach/intervention.

At Level 3, submissions also provide more comprehensive information about the costs of an approach/intervention. Ideally this might include factors such as: the time of anyone involved; resources; and materials used for example. They should preferably also include an initial attempt at cost-benefit analysis, although this is not essential for inclusion at this level. HM Treasury have outlined a methodology for a cost-benefit analysis in their report 'Supporting public service transformation: cost benefit analysis guidance for local partnerships'.

<div style="border: 2px solid #333; border-radius: 20px; background-color: #9a8fbf; padding: 20px;">

Evidence required against the four classification criteria used by the Classification Panel: Level 3

1. **What it is and why it works** - A theory of change/logic model that clearly describes how the approach/intervention achieves its intended outcomes, and/or reference to an existing evidence base for the approach.

2. **The context** - A thorough description of the context in which the approach/intervention operates, including all information about context from Level 1.

3. **Effectiveness** - An evaluation that clearly demonstrates the impact of the approach/intervention. Appropriate methods need to be used which demonstrate that any observed impact are identifiably due to the approach/intervention alone (see text for typical examples).

4. **Cost** - Information about costs required to implement the approach/intervention is presented. A cost-benefit analysis is either presented or considered.

</div>

---

[6] https://www.gov.uk/government/publications/justice-data-lab

Advice and support

If you have any questions about how to use the Standards of Evidence, you can email the YJB Social Research and Effective Practice Team here. The team can also provide advice and support in relation to submitting examples of practice to the library.

Glossary of terms

Comparison time series design - Comparing the changes which occur in one group receiving a treatment over time, with a separate group who did not receive the treatment. The comparison time series group need not be allocated to the intervention specifically and might also be drawn retrospectively.

Control group - A group of people in a study who do not receive the approach or intervention being studied. Instead, they may receive the standard treatment or a dummy treatment. The results for the control group are compared with those for a group receiving the treatment being tested. The aim is to check for any differences. Ideally, the people in the control group should be as similar as possible to those in the treatment group, to make it as easy as possible to detect any effects due to the treatment[7].

Cost-benefit analysis - One of the tools used to carry out an economic evaluation. The costs and benefits are measured using the same monetary units to see whether one exceeds the other[1].

Effective Practice Classification Panel - A panel of YJB representatives and academics working in the area of youth justice. The panel applies the Standards of Evidence to the task of classifying particular submissions in accordance with the quality and extent of the evidence which exists as to their effectiveness. Along with this it is also the task of the panel to: provide brief recommendations to submitters on how to strengthen the design and evaluation of their practice; bring research expertise and independent rigour to this work; and contribute to the review and ongoing improvement of the YJB's Effective Practice Classification System (the panel and the accompanying Standards of Evidence).

Distance travelled - The changes that occur as people progress along a 'pathway' or intervention, from one point to another. Progress is typically measured to assess the level of change that has taken place[8].

Fidelity - Other dimensions of fidelity (other than Programme fidelity) are: exposure (ensuring the appropriate amount/quantity of the intervention is administered and received); reach (ensuring that the approach/intervention meets the needs of the appropriate target group); and engagement (ensuring that users buy into the process).

Logic model - Logic modelling, or programme logic, is an approach for representing the way a programme's various components are expected to fit together to achieve its outcomes. A logic model creates a diagrammatic representation of the key components of a programme and the way that actions are intended to lead to outcomes. The underlying emphasis on logic underlines the potential for a model of this type to challenge assumptions and prompt a reality check about whether or not the programme's interventions are adequate to achieve its intended outcomes[9].

Impact - What is expected to happen long after the intervention has finished[9].

Pre/post design (Also called a 'Before and after study') - An evaluation approach where the outcome measures are measured before and after an intervention has been delivered. The intervention can either be delivered by the investigator or by others

---

[7] https://www.nice.org.uk/Glossary?letter=C
[8] http://www.ces-vol.org.uk/tools-and-resources/outcomes-and-outcome-indicators/measuring-outcomes-and-distance-travelled.html
[9] https://www.england.nhs.uk/wp-content/uploads/2015/06/bcf-user-guide-04.pdf.pdf

(observational before and after study). An approach that is often called a pre–post study[10].

Propensity Score Matching - Propensity scores are an alternative method to estimate the effect of receiving an approach/intervention when randomised allocation is not possible. Propensity score matching refers to pairing individuals who receive an approach/intervention to those who do not. It is then used to compare the outcomes of those two different groups of people[11].

Randomised allocation - Assigning participants in a research study to different groups without taking any similarities or differences between them into account. For example, it could involve using a random numbers table or a computer-generated random sequence. It means that each individual has the same chance of receiving each intervention[12].

Regression Discontinuity design - At its simplest this is a pre/post design. It differs from other pre/post designs however by the method of how participants are assigned to conditions. Participants are assigned to treatment or comparison groups on the basis of a score on a pre-intervention measure. The impact of an intervention is then measured in the same way as any pre/post design, but outcome analysis is undertaken using a more detailed statistical procedure.

Theory of change - The description of a sequence of events that is expected to lead to a particular desired outcome[13]. This is described in more detail within this guide.

---

[10] http://www.nice.org.uk/article/pmg4/chapter/appendix-d-glossary-of-study-designs#before-and-after-ba-studies
[11] http://umanitoba.ca/faculties/health_sciences/medicine/units/community_health_sciences/departmental_units/mchp/protocol/media/propensity_score_matching.pdf
[12] https://www.nice.org.uk/Glossary?letter=R
[13] http://r4d.dfid.gov.uk/pdf/outputs/mis_spc/DFID_ToC_Review_VogelV7.pdf