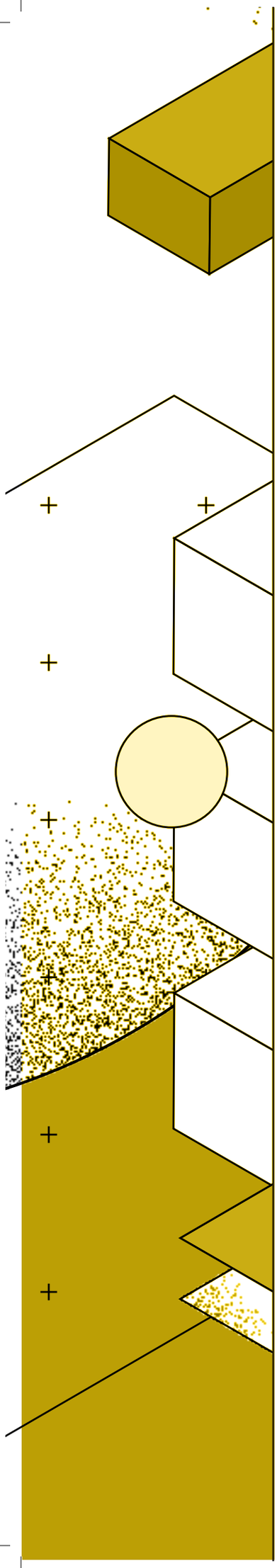


Centre for
Homelessness Impact

Evidence standards and evidence claims in web based research portals

David Gough and Howard White

Campbell
Collaboration

UCL

About the Centre for Homelessness Impact

The Centre for Homelessness Impact champions the creation and use of better evidence for a world without homelessness. Our mission is to improve the lives of those experiencing homelessness by ensuring that policy, practice and funding decisions are underpinned by reliable evidence.

About the authors
Dr David Gough is Professor of Evidence Informed Policy and Practice and the Director of the EPPI-Centre in the Social Science Research Unit at UCL. He came to the unit in 1998, having working at the University of Glasgow and Japan Women's University. Previously David wrote on child protection and abuse, but now spends most of his time on the study of methods for research synthesis and research use.

Dr. Howard White is Chief Executive Officer of the Campbell Collaboration. Previously he was the founding Executive Director of the International Initiative for Impact Evaluation (3ie) and before that led the impact evaluation programme of the World Bank's Independent Evaluation Group.

Acknowledgements
Thanks to all the fourteen organisations for their evidence portals that provide the crucial service of access to and information on research evidence in their fields. Thanks also to the many portals that provided very helpful comments and feedback on our summary of their portals. Finally, thanks to CHI for their support and for funding this study.

Centre for Homelessness Impact
www.homelessnessimpact.org
Crisis
Registered Charity Numbers:
E&W1082947, SC040094.
Company Number: 4024938

This report should be referenced as: David Gough, Howard White (2018)
Evidence standards and evidence claims in web based research portals.
London: Centre for Homelessness Impact.
Available from: www.homelessnessimpact.org

ISBN: 978-1-9995928-3-7

Evidence standards and evidence claims in web based research portals

David Gough and Howard White

Foreword

Our journey to create a Centre dedicated to championing the creation and use of reliable evidence started with conversations; conversations with hundreds of people working towards a future without homelessness in the UK and overseas.

We also talked to individuals and organisations with similar missions in other social policy fields and learned many valuable insights from them.

These interactions shaped our initial programme of work and led us to prioritise the release of an Evidence Portal, our 'Intervention Tool', when we launched the Centre in the Spring.

Why?

In short, we heard from our 'end users' that collectively we are not doing enough with the evidence that already exists, and that its effective use is more often than not hampered by the sheer volume and confusing array of evidence available.

Another common challenge cited was understanding the reliability of different evidence claims.

We looked outside our field for inspiration and found many 'best in class' examples of tools created in response to similar challenges. We found there was appetite for an evidence portal that could help commissioners and funders of homelessness services and others make better use of limited resources by providing an accessible and reliable entry-point to the body of knowledge on the impact of interventions without having to sift through lots of studies.

The Education Endowment Foundation's Toolkit in particular appealed to people and seemed to be working well so served as our model as we developed our Intervention Tool. At a glance, it provides one-page overviews for each of the main interventions. The overviews give ratings for how reliable the body of evidence is, how cost effective the intervention is and what kind of impact it has for people experiencing homelessness.

We knew that to 'work' the Intervention Tool would need to be trusted. This is vital if insights are to be acted upon. So a crucial first step was to develop the evidence standards that would inform our evidence portal - i.e. a clear, rigorous and transparent account for the approach taken to assessing the quality and reliability of the evidence in the Intervention Tool.

We invited David Gough and Howard White to help us with this task because of their expertise in this area and a sensible way to approach the task seemed to be to review the standards of evidence used by the best well known evidence portals from across the world.

Throughout our journey we have been able to move fast because of how much we've learned from those with a similar mission to us in related fields who have come before us and the development of our evidence portal was no exception. We believe this is a good example of how it is possible to combine rigour and agility - a must if knowledge broker organisations like ours are to produce work that is both trusted and timely.

We welcome the recommendations in the report and we are committed to adopting them as we continue to develop the Intervention Tool over the years to come. We look forward to continuing to work with our academic partners, user experience designers and our end users to further build the Intervention Tool.

It may indeed be the case that a universal standard may not be possible but there is nevertheless a lot to be said for not reinventing the wheel. We hope this short study will be useful to others developing evidence portals in other social policy fields.

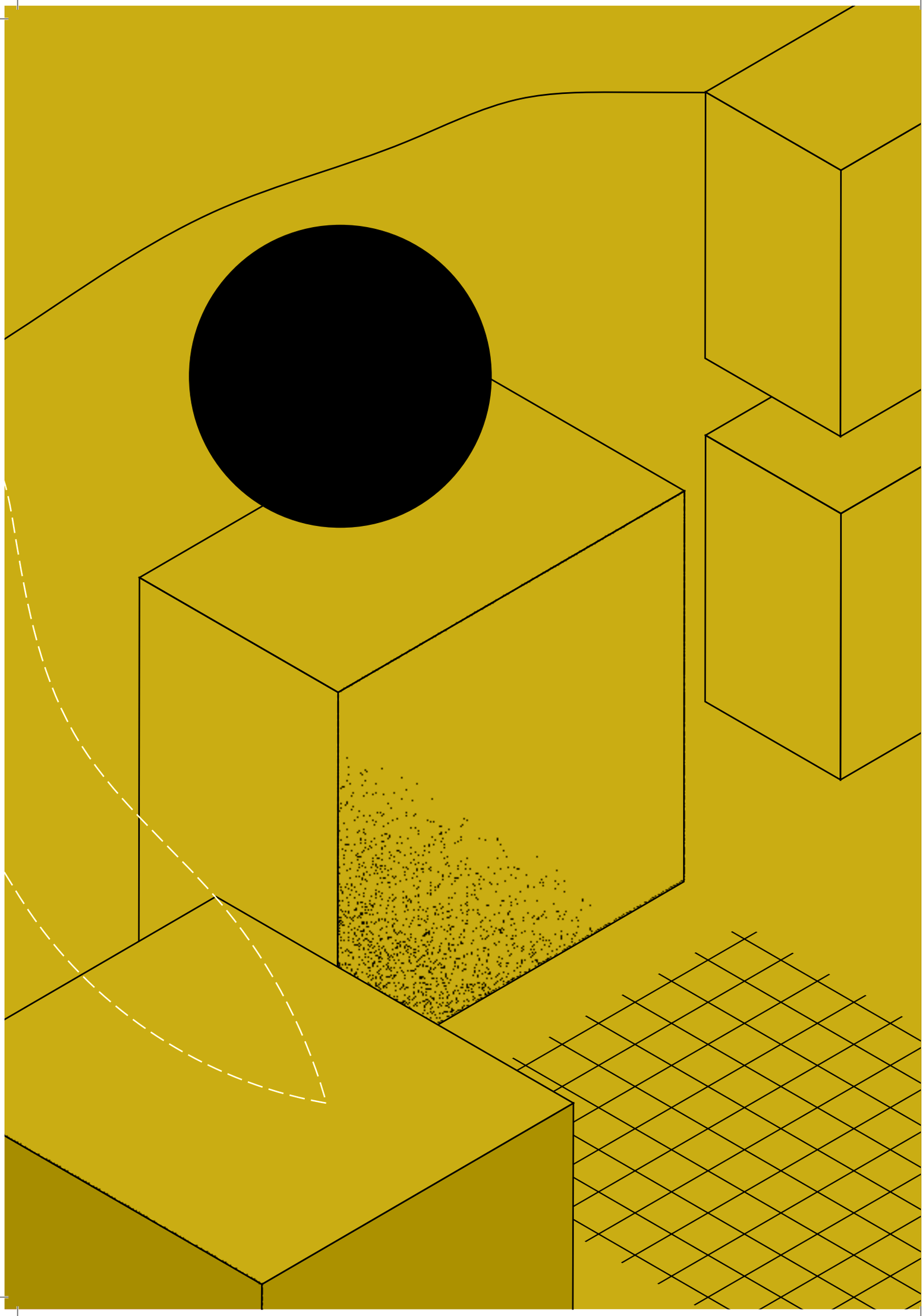
Dr Ligia Teixeira
Director, Centre for Homelessness Impact

Contents

Chapter 4

References

Appendices



Executive summary

A survey was undertaken for the Centre for Homelessness Impact of 14 English language evidence portals to examine the type of research evidence that they contained and the evidence standards used. It was found that portals predominantly focused on questions on the efficacy of interventions using experimentally controlled studies. Evidence standards were based predominantly on the trustworthiness of individual primary research studies and of the whole of the existing evidence base. The report makes the following recommendations about the use of evidence standards in evidence portals:

1: Specify the aims and methods of making evidence claims and ensure consistency across levels of evidence (guidance, evidence base, included studies).

2: Consider using broad rather than narrow evidence base questions.

3: Use explicit rigorous methods of evidence synthesis to make claims about the existing evidence base.

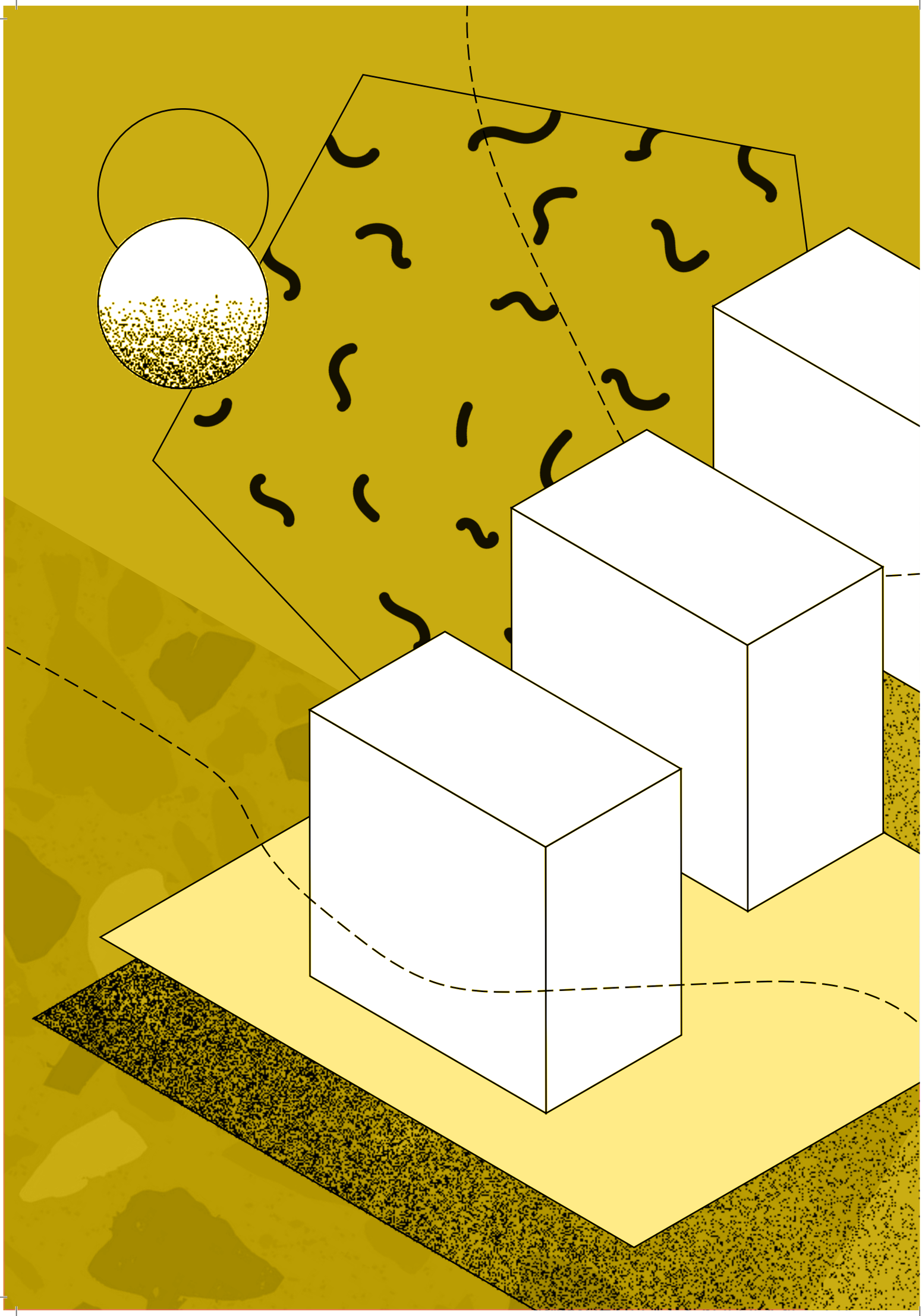
4: Specify and justify the different evidence standards for making different claims about the existing evidence base including impact, strength, extent and consistency of evidence, process and contexts and costs.

5: Specify and justify evidence standards for included studies.

6: Specify methods as well as criteria for achieving evidence standards.

7: Develop methods and standards for policy and practice guidance informed by the evidence base.

These recommendations are considered in relation to CHI's plans to develop their evidence portal.



Chapter 1

Introduction

1.1 Background and Aims

1.1.1 Background

Evidence informed decision making is simply the use of research evidence as part of the decision making processes. Policy, practice and personal decisions are influenced by numerous conscious and unconscious factors. These factors include values, resources, information on the current and potential future situation and the likely effect of different decisions being made. To the extent that information is one component of decision making, then why not ensure that the information is based on the most reliable research available?

> Research simply means enquiry based on fit for purpose, rigorous, and explicit methods. There are many such methods. Evidence informed decision-making should not be the sole domain of one type of research evidence.
>
> This is not an argument that decisions should be made on the basis of research evidence alone. It is that evidence from reliable research is a potentially useful source of information and should be made available in ways it can best help inform decisions.
>
> Despite the appeal of the argument that research evidence provides useful information for decision making, most decisions are made without the use of such research evidence. There are a variety of reasons for this. It may be that relevant research evidence is not available. Even if research is available it may not be of sufficient quality to be relied upon. It may be available and trustworthy, but the decision makers may not be aware of the research or have the motivation or opportunity to use it (Langer et al 2016). Research outputs are often inaccessible both in terms of their style and

content, and also because of pay walls and low discoverability. Busy policy-makers and practitioners do not usually have time to read academic papers even if they are discoverable and accessible. And if they do read them, the implications of the findings for policy may not be clear or different papers may contradict one another. In other words, the assumed evidence ecosystem linking research use with research production could operate better. It is for this reason that recent years have seen the emergence of knowledge brokering as link between research and its use.

The development of knowledge brokering as a distinct activity is manifested in the creation of individual knowledge broker roles as well as intermediary organisations to help improve the links between research and its use. The influence can be two way. Those making or affected by decisions can influence the research that is undertaken and research can be more accessible and more relevant to potential users of that research.

The research evidence that could be considered by such brokerage organisations is immense. Even within one topic area there are many questions to be asked and many research methods to address them. In practice, a lot of these organisations focus on questions of efficacy or effectiveness (the 'What Works' movement). What Works organisations mainly focus on whether the intervention (action) is likely to achieve the desired and intended effect or impact.

The research on the efficacy and effectiveness of different strategies for increasing the use of research in decision making has been reviewed by Langer and colleagues (2016) using a framework of six mechanisms and three behavioural components of engagement between research users and research producers. In practice, brokers and brokerage organisations undertake a wide range of activities (see for example, Gough et al 2018). One common approach is direct (person to person) interaction where a broker discusses evidence needs with policy makers and then provides them with an overview of the existing evidence, often helping them interpret the implications for their own policy problem. This is one of the approaches taken by research analysts within government. Specialist brokerage centres, such as, for example, government-

funded research centres in the Nordic countries produce policy-oriented evidence synthesis for health, education and social welfare for topics agreed with government departments.

Such direct interaction can work well when providing a brokering role for a single decision-making agency such as a central government department. But it is less well suited when decision-making is decentralised with decisions being taken by prison governors, school managers, case workers or NGO programme managers. In such cases developing evidence resources that more generically summarise what is known from research such as evidence portals, guidance and checklists may be a particularly helpful resource.

These types of evidence product have the common element of not requiring the user to consult the underlying studies. They differ in the amount of curation by the broker and agency left for the decision maker. Evidence portals mostly aim to make the evidence available to support an evidence-informed decision by the decision-maker. Guidelines and guidance, prepared by expert committees on the basis of evidence reviews, make recommendations on the decisions to be taken. And checklists, produced using a similar process to guidelines, suggest what practitioners could or should do.

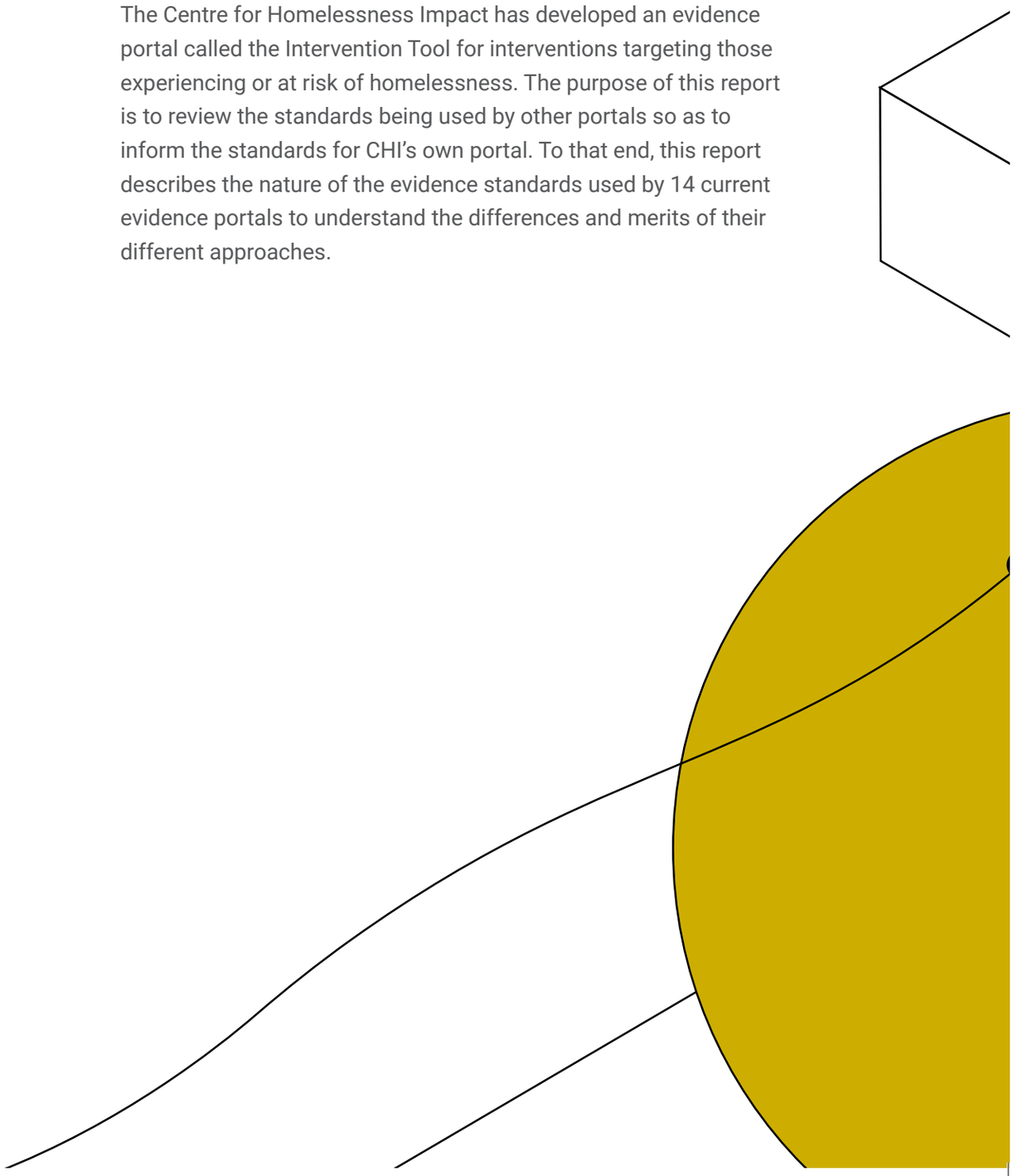
Portals have a range of formats. Some are more like a database or repository of resources. We would call those evidence hubs rather than portals. Our interest is in the more heavily curated portals which provide web based information that summarises what is known from research about different issues. They are called evidence portals as they act as a doorway to accessing research findings. They are also called clearinghouses or toolkits. They are different to evidence hubs which are networks for sharing information.

Users of research evidence need to know how strong, trustworthy and relevant any research findings are. They need to know what standards are being applied to make a judgment about the findings of research. They need to know the basis for the evidence claims being made. Without this they will not know whether they should use the research findings to inform their decisions. Therefore for evidence portals to be useful and trustworthy and to advance the

13

evidence informed policy and practice, they need to have explicit coherent evidence standards. By evidence standards we mean the approach taken to developing and assessing the quality of research and the soundness of the research findings presented in the portal. It is important to have clear, sound and transparent evidence standards to underpin the credibility of the evidence portal. They provide the basis for making evidence claims (and protect the organisation from criticisms of bias or partiality in their presentation of the evidence).

The Centre for Homelessness Impact has developed an evidence portal called the Intervention Tool for interventions targeting those experiencing or at risk of homelessness. The purpose of this report is to review the standards being used by other portals so as to inform the standards for CHI's own portal. To that end, this report describes the nature of the evidence standards used by 14 current evidence portals to understand the differences and merits of their different approaches.



1.1.2 Study Aims

The first aim of the study is to undertake a survey and analysis of 14 current research portals to examine the nature of the resources provided and the evidence standards applied.

The study does not aim to provide a comprehensive survey of evidence portals. It uses a purposeful sample to examine a range of well-known English language based portals in the USA and Europe. It is a small scale study exploring the nature of a phenomenon rather than attempting to describe the prevalence of different features.

The study also does not aim to provide an examination of the organisations that provide the portals. In some cases, an organisation's work is predominantly their portal but other organisations have broader remits and activities.

The second aim of the study is to propose recommended practice on the evidence resources and evidence standards used in research portals.

These first two aims support the third and related aim: to inform the development of the Centre for Homelessness Impact Intervention Toolkit.

1.2 Study methods

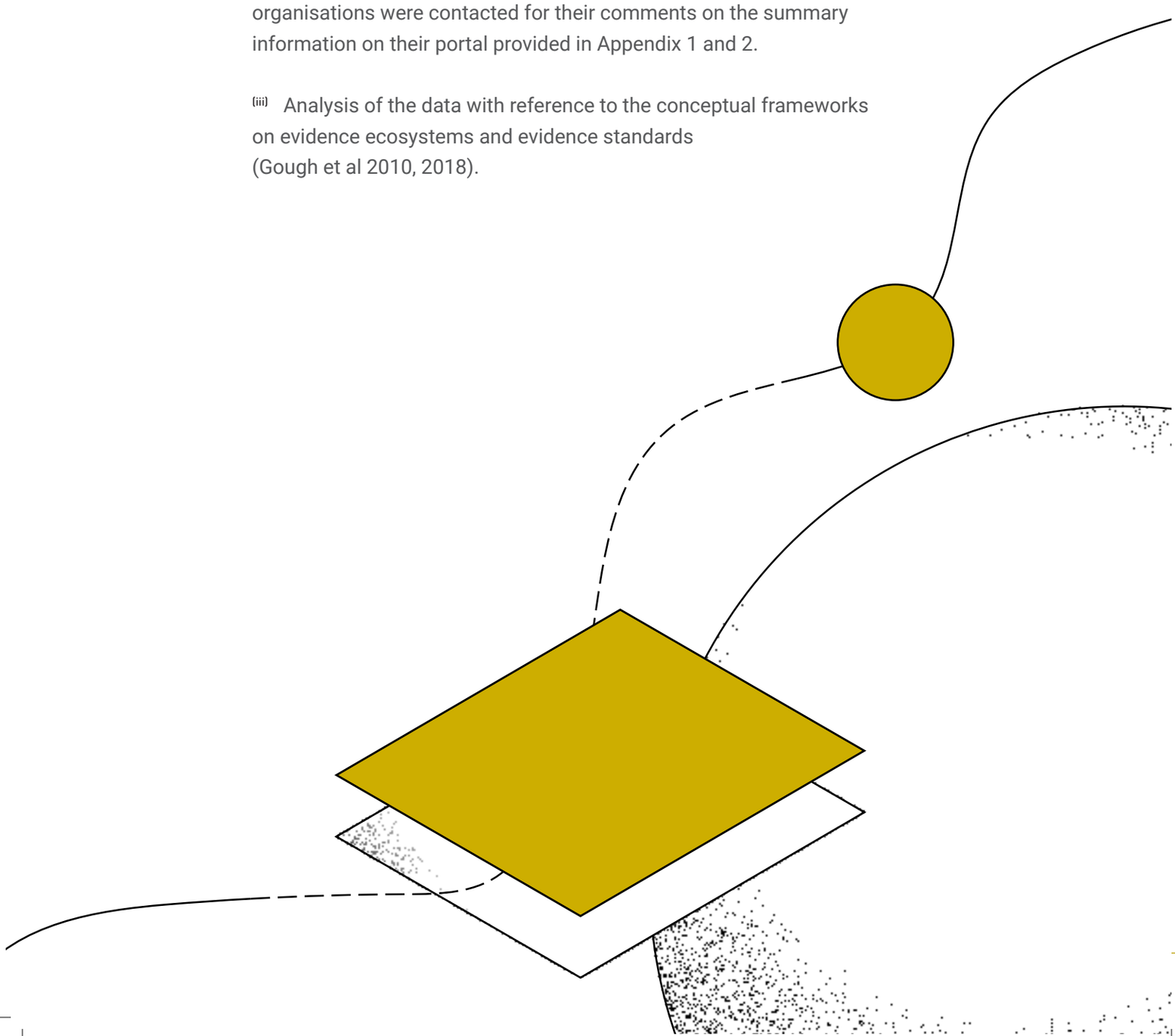
[i] Purposive sampling of 14 evidence portals based on the research team's knowledge of English language research portals in various areas of social policy concerned with the evidence on the effectiveness of certain courses of action (interventions).

- California Evidence-Based Clearinghouse for Child Welfare (CEBC), USA
- Clearinghouse for Labor Evaluation and Research (CLEAR), USA
- Conservation Evidence (CE), UK
- Early Intervention Foundation (EIF) Guidebook, UK
- Education Endowment Foundation (EEF) Toolkit, UK
- European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) Practice Portal Evidence Database, EU (based in Portugal)
- European Platform for Investing in Children (EPIC): Evidence Based Practices, EU (managed from UK)
- Evidence Based Teen Pregnancy Programs (EBTP), USA
- Institute of Educational Sciences (IES) What Works Clearing House, USA
- National Institute for Health and Care Excellence (NICE), UK
- What Works Centre for Crime Reduction: Crime Reduction Toolkit (WWCR), UK
- What Works Centre for Local Economic Growth Toolkits (WW G), UK

- What Works in Reentry Clearing House (WW RCH), USA
- What Works Wellbeing Evidence Comparison Tool (WW Wellbeing), UK

(ii) Survey of the content of the portals and their reference to their methods including manuals and guides. This is based on the materials provided by the portals on their websites and does not refer to further materials or other information from the host organisations. Structured data collection was undertaken by David Gough, with a selection checked by Howard White. All portal organisations were contacted for their comments on the summary information on their portal provided in Appendix 1 and 2.

(iii) Analysis of the data with reference to the conceptual frameworks on evidence ecosystems and evidence standards (Gough et al 2010, 2018).



Chapter 2

Evidence resources, evidence standards and evidence claims

This section of the report explains the analytical distinctions and framework used in this study.

Evidence portals vary in both the types of resource being offered and the evidence standards used as a basis for the evidence claims being made. These two dimensions – (1) types of resources, and (2) the evidence standards - form the main basis of the analysis of the portals in this survey and are discussed in turn in Sections 2.1 and 2.2.

2.1 Nature of evidence claims in different parts of an evidence ecosystem

The first dimension concerns the nature of the resources and evidence claims being offered. A portal may provide access to evidence from: (i) the findings of individual studies; (ii) a broader statement of the evidence base from a synthesis of findings across a number of studies; or (iii) evidence informed guidance and recommendations on what decisions to make based on the interpretation of the findings of research studies together with other contextual information.

These different types of resource - individual studies, synthesis of evidence base, and evidence informed guidance/ recommendations - are clearly related to each other. A research portal providing information on the evidence claims about individual studies will be doing so in order to inform the evidence base about such interventions. A portal which provides information on an evidence base on the effectiveness of an intervention has the implicit

or explicit aim to inform judgements about whether such an intervention should be applied in practice.

The three different types of resources are shown schematically as part of an evidence ecosystem as shown in figure 1 (see also, Gough et al 2018). On the left hand side of the figure is decision making. On the right hand side is the production of research. For research to inform decision making there needs to be some form of engagement between the use and production of research.

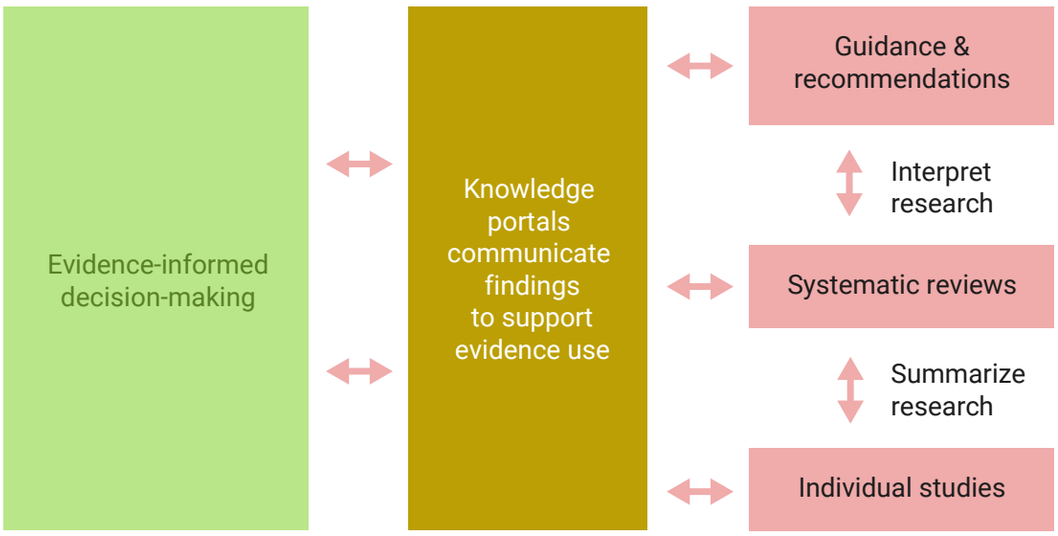


Figure 1. The position of evidence portals in the evidence ecosystem

The nature of each of the three types of evidence resources is now discussed in turn.

Individual studies

Individual studies are studies of a particular intervention or set of interventions. The studies may be quantitative analysis of effects, or qualitative studies of issues such as context, beneficiary perspectives and implementation issues.

Decision makers can engage directly with the findings of individual primary research studies. But there are several challenges to doing so.

The research may be:

- Difficult to access.
- Difficult to understand.
- Not trustworthy due to methodological quality (low internal validity) or the extent of the evidence on which conclusions are made.
- Trustworthy but not representative of other research findings (i.e. the wider evidence base).
- Trustworthy and representative but not relevant to the specific focus and context of the decision being made.

Even if these challenges are overcome, the process of accessing, understanding and appraising a research field – many of which are growing rapidly - is a time consuming and skilled activity. Decision makers such as policy makers, professional practitioners and members of the public may not have such skills and time. For this reason it is likely to be more efficient for them to use overviews of the evidence base for the issue of interest.

An evidence base (evidence synthesis)

Statements about an evidence base go beyond one individual study to make a statement about what is known (and thus also not known) from research in relation to a research question. The ways that statements are made about 'what is known' can vary. They can be based on:

(i) An explicit research process for synthesising the research evidence on a particular question (a systematic review). A systematic review is a review of research evidence using explicit rigorous methods. It is a piece of research but instead of undertaking new primary research, it addresses a research question by examining the findings of already existing research studies.

(ii) A more informal and implicit process for synthesising

the research evidence on a particular question (a non-systematic review).

(iii) An assessment of one or more rigorous research studies identified as reporting a particular finding (statement of sufficient evidence without necessarily checking the full evidence base).

Guidance and recommendations

Evidence-based guidelines and recommendations are produced based on a review of the evidence. Their purpose is very clearly to guide decision makers based on research evidence and other factors. This guidance is based on an expert interpretation of research (often in consultation with other stakeholders) on the basis of:

(i) Various aspects of the particular context in which the target group of decision makers operate, as this context may not be the same as in that in which the studies in the evidence base were conducted.

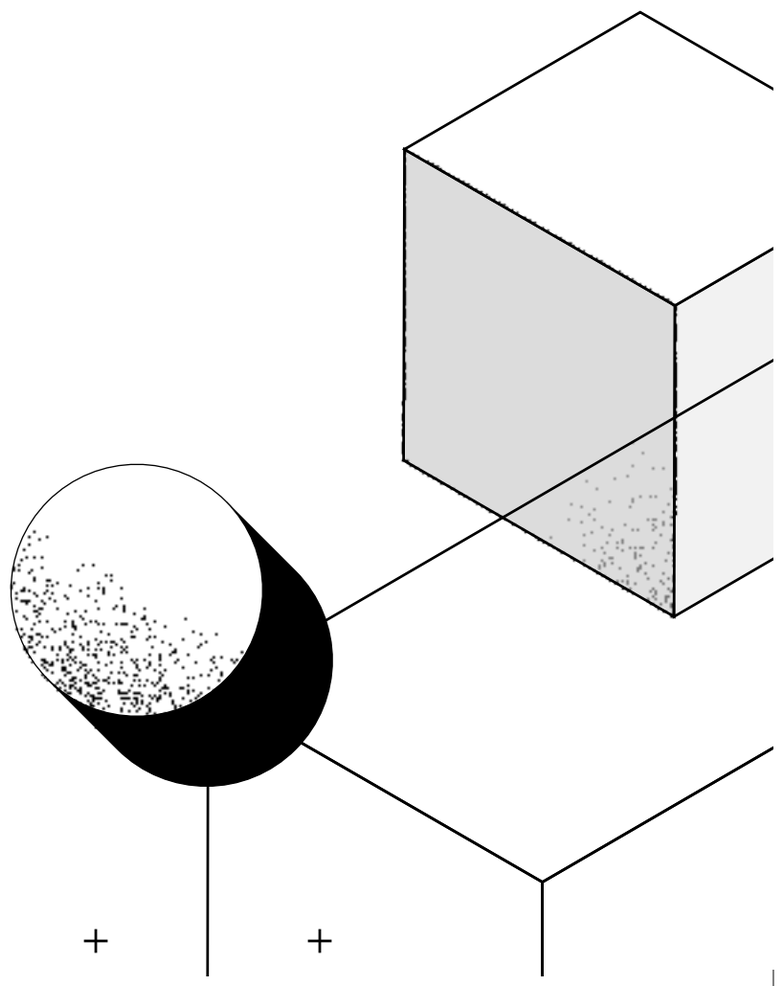
(ii) The particular perspectives (values and priorities) of the different stakeholders involved in decision making or those affected by the decision.

Individual decision makers can engage directly with an evidence base but just as with engaging with individual studies this requires skills and time. Also, a decision maker may not have access to or be able to sufficiently engage with the various stakeholder views. There are therefore efficiencies and strengths from interpretations of an evidence base within particular contexts and perspectives being undertaken by a local or national group of experts, users of services and decision makers, presenting their conclusions in guidance and recommendations.

Nature of efficacy and impact questions

Another issue is the extent that a research question is framed on a particular problem or a particular intervention. For decision makers, the starting point is often an issue or problem for which they require a solution. Research may help to identify the most effective solutions.

However, decision makers may be presented with certain interventions as solutions and so they then seek evidence about the efficacy of those solutions. This approach is common with pharmaceutical products and branded intervention programmes in social policy where the 'actions' are proposed and then evidence about their effectiveness is sought. This evidence may come from individual studies or evidence synthesis. The 'search for evidence' to support specific programmes may encourage the 'one or two good studies' approach to an evidence base where an intervention programme is determined to be effective as some rigorous studies report positive findings. As argued below, making decisions on the basis of one or two studies, however trustworthy, does not provide a rigorous assessment of the potentially broader evidence base.



2.2 Evidence standards for making evidence claims

The second dimension concerns the evidence standards which are the basis of the evidence claims provided by the portal (whether that claim is based on individual studies, a review of the evidence base or guidance/ recommendations).

2.2.1 Information needed from and about the evidence base

The main focus of this study are portals which provide information on which interventions 'work'. That is whether specific approaches or programmes are effective in achieving certain outcomes. For example, the Education Endowment Foundation's Teacher and Learning Toolkit reports the effectiveness of 34 interventions such as 'arts participation', 'feedback' and 'repeating a year'. In contrast the Institute of Education Sciences What Works Clearing House structures the evidence by branded programmes such as 'Earobics', 'DairyQuest' and 'SpellRead'. Similarly, Evidence-Based Teenage Prevention lists branded programmes such as 'Be Proud, Be Responsible, Be Protected' and 'Seventeen Days'. This focus on branded programmes in the US portals reflect the nature of social service and education provision in that country.

> There are several different things that users of research evidence might want to know about the evidence base for the effectiveness of interventions to respond to social issues. They may, for example, be interested in the strength of the evidence, the process by which any impact is achieved, the context within which the evidence has been found, and the costs or cost effectiveness of any impact.
>
> (i) Effectiveness: What is the size of the impact of an intervention?
>
> (ii) Strength: What is the strength and extent of evidence for that being the extent of impact?

(iii) Process: What is the process or critical components of an intervention relevant to decision making about undertaking interventions in different contexts (e.g. implementation issues)?

(iv) Context: Where is the evidence from? This allows users to judge the relevance of the evidence to their context.

(v) Cost: What are the resources or costs required to implement the intervention?

2.2.2 Evidence standards for different components involved in an evidence base

All of the evidence claim types listed in section 2.2.1 are based on different types of evidence. They therefore require different research methods and different standards for making such evidence claims. What are the criteria used to determine, for example, how strong the evidence is for effectiveness? How are judgements made as to whether these evidence claims are justified and how these criteria are applied (Gough 2016, Liabo et al 2017).

Evidence standards are an approach to developing and assessing the quality of research and the soundness of the research findings, which capture not only internal validity (internal methodological coherence of a research study, i.e. if the results are credible) but also external validity (the extent that what the research measures really is what you wish to measure in the real external world and so the results can be applied elsewhere).

Clearly there can be no one universal evidence standard since the standards differ according to the research question, the type of research product (single study, evidence synthesis, or guidelines) and the research design (e.g. quantitative or qualitative). There are many evidence standards available for assessing the quality of single studies. Some standards are presented as checklists of criteria. The Critical Appraisal Checklist site maintained by Cardiff University Specialist Unit for Review Evidence provides several evidence standards classified by study design, e.g. experimental

studies, observational and qualitative studies. Similarly there are standards to assess systematic reviews such as AMSTAR 2 and SURE. There are also existing standards for assessing bodies of evidence, of which the best known is GRADE, which can also be applied to guidelines. Agencies adopting an evidence-based approach may also issue evidence standards. The UK Department for International Development published 'Assessing the Strength of Evidence' in 2014 covering the assessment of both single studies and bodies of evidence (DFID, 2014).

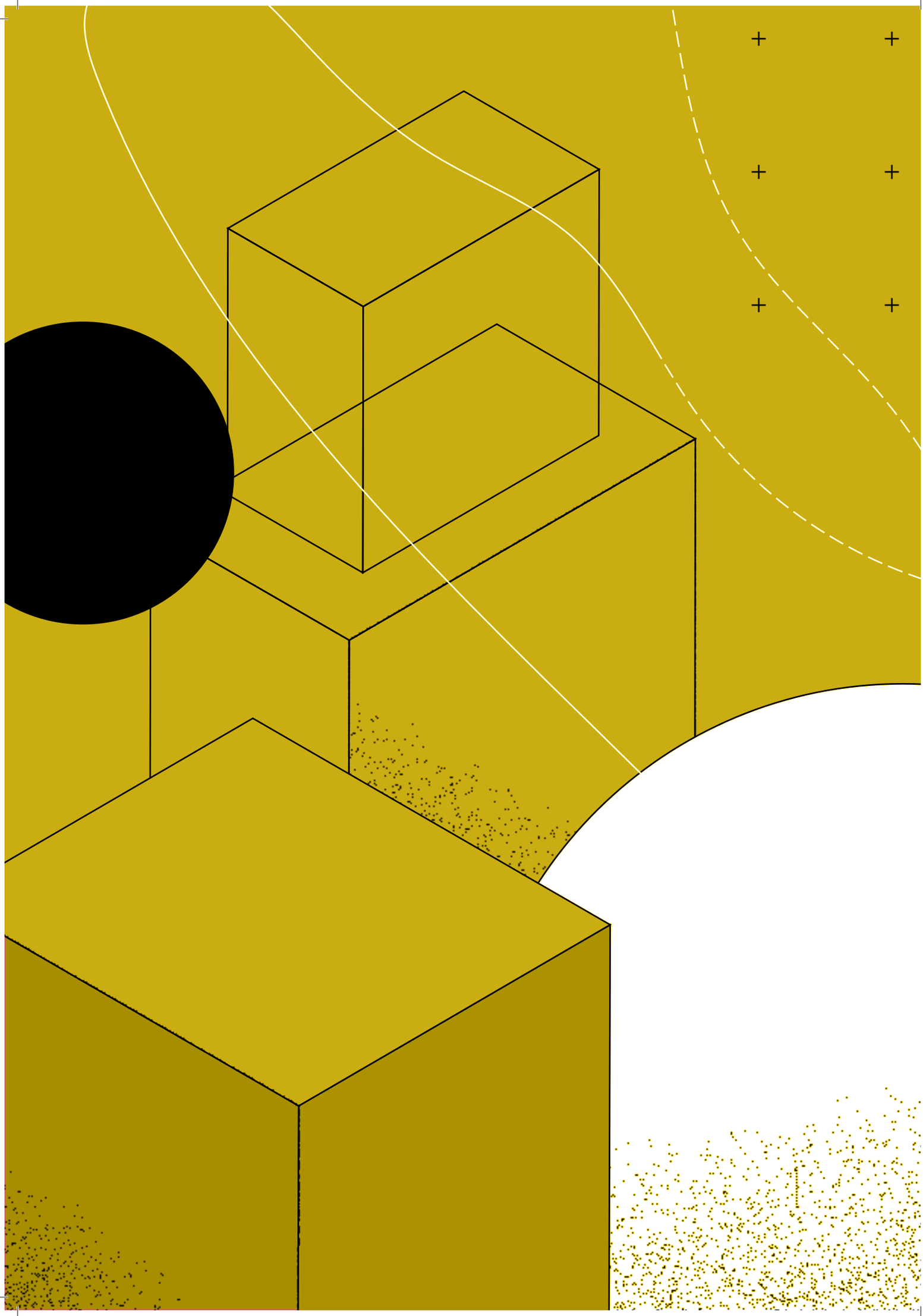
Many evidence standards are concerned with technical quality, but, quality appraisal also includes fitness for purpose of the method and the relevance of the focus of a study for answering a research question. So the synthesis of an evidence base involves multiple evidence standard components (Gough 2016):

(i) Evidence standards for appraising the methods for undertaking a review of the evidence to make a claim about an evidence base. This includes: (a) the technical quality of the evidence base review; (b) the fitness for purpose of the method for the review question; (c) the relevance of the focus of the review for the review question.

(ii) Evidence standards for appraising the quality and relevance of individual studies included in a review of an evidence base. This includes: (a) the technical quality of the included study; (b) the fitness for purpose of that method for the review question; (c) the relevance of the focus of the study for the review question.

(iii) Evidence standards for appraising the totality of evidence included in a review of an evidence base. This includes: (a) the nature of the totality of evidence; (b) the extent and distribution of that evidence.

(iv) Evidence standards for appraising different evidence claims made by a review. This depends upon the type of claim being made and can be based on a combined appraisal of (i), (ii) and (iii) above (and these components might be given different weights in making such a combined appraisal).



Chapter 3

Study Findings

A structured description of basic characteristics of each of the 14 portal platforms is given in Appendix 1. A structured description of each of the portals' methods and evidence standards is given in Appendix 2. These appendices provide the data for the description of the results and the following discussion and recommendations.

3.1 Questions, methods and evidence standards for studies included in an evidence base

Research questions

One way in which the portals differed was in the type of research questions that they addressed. All of them were concerned with the effectiveness of certain interventions, but they varied in how these were framed. As mentioned above, most of them asked questions about the effectiveness of branded programmes. The California Evidence-based Clearing House for Child Welfare (CEBC), for example, provides an assessment of over 200 branded programmes which affect child welfare. A few portals take a broader approach in examining the effectiveness of an approach rather than a particular programme. The Education Endowment Foundation's Teacher and Learning Toolkit, for example, presents evidence of the effectiveness of 34 types of intervention such as arts participation and repeating a year in school. And some examine both programmes and broader approaches. One portal, NICE, is focused on producing practice guidance and they take an even broader approach in starting with a practice issue – such as diet, physical activity or child abuse - and then considering what interventions might help and what evidence existed for the effectiveness of those interventions.

27

Review methods

Another way that the portals varied was in their method for summarising an evidence base.

One approach is to use existing systematic reviews. Where possible, EEF attempts a statistical meta analysis from previous reviews of the education approaches presented in the portal. WW Crime instead uses the strongest components of the most rigorous and relevant reviews to summarise the state of knowledge. EMCCDA uses prior reviews to provide a narrative review of an evidence base.

Another approach is to undertake new systematic reviews, used for example by EPIC, NICE and WW Wellbeing. Other portals undertake reviews which are less obviously systematic. Both CLEAR and WW Growth use specific scales to assess primary studies with narrative synthesis of studies meeting a threshold for methodological quality. CE conducts systematic searches for literature and then uses an expert model of creating synopses of the evidence base.

A different approach is taken by CEBC, EIF, and EBTP. Instead of reviewing the whole of an evidence base, they make judgements about the effectiveness of programmes on the basis of there being at least one or two rigorous studies with evidence of beneficial effect. The WW RCH portal is similar in providing access to individual studies rated for strength of evidence to support their conclusions. IES also makes judgements on the basis of two rigorous studies as long as there are not also studies indicating overriding contrary evidence.

Included studies study design inclusion criteria

All of the portals are concerned with the effectiveness of interventions and so, in assessing an evidence base, they seek out research studies that are evaluations of effectiveness.

Research designs differ in how powerful they are at appraising the effectiveness of an intervention. Some portals restrict the type of research designs that they will consider. They may only allow rigorously undertaken powerful experimental designs such as randomized controlled trials (RCTs). On the other hand, they may allow all forms of RCT as well as non-experimentally controlled designs.

The extent that evidence portals specify certain research designs varies. CLEAR is unique amongst the 13 portals examined in only considering evidence from RCTs. In contrast, CE's inclusion criterion of studies which have 'quantitatively monitored the effect' allows inclusion of before versus after analysis with no comparison group. EIC also includes studies with pre-post designs without a comparison group. The other portals either base their analysis on reviews, or include both experimental and non-experimental designs with a comparison group.

Less common factors include WW RCH requiring that the researcher be independent of the intervention unless the study has been peer reviewed. The first criteria aims to limit bias that may occur from investigator belief in the effectiveness of an intervention yet this rule can be avoided by publishing in a peer reviewed output which is not a very strict criteria. This is in contrast to other portals presenting branded programmes – such as CEBC – in which the supporting studies are frequently conducted by the programme designers with no explicit mention or consideration of the conflict of interest in the portal. CE is the only one of the portals examined to overtly consider non English language journals.

Included studies appraisal (evidence standards)

Even if evidence portals are not strict about what research designs may be considered, they can still appraise such included studies for methodological quality and relevance.

> Most of the portals include studies that seek evidence of causal effect whilst controlling for extraneous variables (the counterfactual), but they vary in the explicitness and strictness with which this is done. Portals such as IES, NICE, WW Growth have very detailed methods and/or criteria for appraising causal inference.
>
> The standards are usually applied using a scale. What Works in Reentry has just two points on its scale (basic and high). IES has a 6 point quality criteria scale and then grades studies according to: meets standards, meets standards with reservations, and does not meet standards. Assignment of a study on a scale depends on a number of study design features, the most notable one being experimental or non-experimental. RCTs are commonly the only design which can receive the highest quality rating: this is the case for CLEAR, IES and EBPT in the US and the What Works Centre for Local Economic Growth in the UK. Other factors are also taken into account, such as attrition and the outcome measure used. EIF's assessment includes 33 criteria.

3.2 Evidence bases: data and evidence standards

Extent of impact / Strength of evidence

Most of the portals use scales to describe the strength of evidence of a beneficial causal effect (and the avoidance of harms) whether this be on the basis of a review of reviews, a statistical meta-analysis of systematic reviews, a systematic narrative review, a non-systematic review, or judgments made on the basis of one or two rigorous studies showing an effect. This can include specification of the consistency of the evidence (for example by IES) as well as the extent of evidence (as in portals using GRADE).

Some of the portals refer to evidence of enduring outcome (EPIC) or transferability of findings shown by evidence from more than one context (EPIC). Others refer to the extent of the evidence in terms of numbers of studies, total sample size of included studies or consistency of findings across studies. Sometimes this is part of the narrative description of findings across studies. Sometimes it is considered in scales of effectiveness with ratings such as 'mixed effects'. Sometimes it is part of a formal statement about the nature and extent of the evidence as in CE scales and in a different way by the GRADE framework used, for example, by NICE.

The California Clearinghouse has a five point scale: 1. Well-Supported by Research Evidence, 2. Supported by Research Evidence, 3. Promising Research Evidence, 4. Evidence Fails to Demonstrate Effect, and 5. Concerning Practice (for example, harmful effects).

EEF's Teacher and Learner Toolkit displays the strength of evidence using between one and five lock symbols. These range from one lock ('Very limited: Quantitative evidence of impact from single studies, but with effect size data reported or calculable. No systematic reviews with quantitative data or meta-analyses located') to five locks ('Very Extensive: Consistent high quality evidence from at least five robust and recent meta-analyses where the majority of the included studies have good ecological validity and where the outcome measures include curriculum measures or standardised tests in school subject areas').

Process by which effect occurs and the context of such effects

The portals may describe potential causal mechanisms in narrative text. WW Crime, however, has the EMMIE framework to systematically seek and rate, if available, mechanism, moderator and implementation data from systematic reviews. There is an overlap between process data and implementation with portals such as EBTP developing logic models for implementation, and so flagging issues which may arise in implementation. But some portals – such as both of the education portals – do not include such information, though both EEF and IES do produce Guidance documents which address implementation issues (see below).

- Whilst not directly related to process, some portals do allow the evidence to be filtered by context. The IES Clearing House has filters to restrict the evidence to specific types of school, grade and so on.

Costs of interventions

Most portals do not have cost information. Some have descriptive information about what resources are needed to provide an action (CEBC, EBTP), others provide a scaled estimate of the costs of such provision (EIF, EEF, WW CR, WW Growth). For example, EEF shows cost using between one and five £ symbols. The lowest cost (£) is up to £2,000 per year per class of 25 pupils, or less than £80 per pupil per year, and the highest (£££££) is over £30,000 per year per class of 25 pupils, or over £1,200 per pupil.

- NICE models cost benefit analyses to make Quality Adjusted Life years (QALY) calculations that are used to determine resource allocation in the English National Health Service.

Methods, processes and criteria

A final issue is the transparency of the ways in which these methodological and evidence standards processes are specified. Some portals have their own or refer to other manuals that specify the methods that need to be followed to ensure that rigour is maintained (CEBC, CLEAR, EBTP, IES, NICE, WW CR, WW G and WW Wellbeing). Others focus more on the procedural steps of undertaking the methods or the criteria for appraising the standard of evidence achieved (CE, EIF, EPIC and WW RCH).

3.3 Practice guidance

In addition to appraising evidence bases related to different interventions, some portals also provide guidance and recommendations on how to act. In some cases, this is simply advising on issues to consider if implementing an intervention (CLEAR, EMCDDA and WW Growth). In other cases, it is more specific advice on the implementation process (CEBC). IES, EEF and NICE use panels and committees to create formal guidance and practice guides informed by evidence and stakeholder values beyond the evidence of efficacy. IES has a check list for information that is used to inform practice guides.

> Preparing guidance is more than simply summarising the evidence. It assesses the evidence, taking into account both applicability and value issues which may arise in applying the evidence in a specific context. This approach is most developed in the health sector: NICE has a social values policy, public consultation and an elaborate stakeholder guidance committee process.
>
> Finally, this chapter on findings should note that the portals tended to explain the particular approach that they took to evidence standards rather than locating their decisions within the many possible components and complexities of evidence standards referred to in Chapter 2.

Chapter 4

Discussion and Recommendations

1. Portal theories of change and justification and balance in methods and evidence standards

- The portals have different histories, resources, audiences and contexts and so it not surprising that they differ in form and function. The approaches to evidence standards also vary. There may be good reasons for this though this is not always clear as the reasons are not always well explained. Similarly, it is not always clear why some portals devote more attention in both detail of explanation and complexity to different parts of the evidence ecosystem that they are covering. Some portals, for example, have very detailed methods or criteria for appraising included primary studies but then provide relatively little detail on how the evidence base as whole is to be appraised. All evidence systems are as strong as their weakest link, which is an argument for consistency in the detail and complexity of evidence standards in primary research and reviews of that research (so in different parts of the evidence ecosystem as in Figure 1).

Recommendation 1

Specify the aims and methods of making evidence claims and ensure consistency across levels of evidence (guidance, evidence base, included studies)

(i) Consider the intended audience and use of the portal and so consider and justify the portal's aims and methods, the specific role of research and guidance production, and how the portal would know if it was achieving its mission.

(ii) Consider the consistency in both detail and complexity of the evidence standards in different parts of the evidence ecosystem covered by the portal, i.e. is the evidence portal fit for it's purposes?

2. Research questions

The main purpose of the evidence portals is to enable engagement between the use and the production of research to inform policy, practice and personal decision making.

The portals are mainly concerned with the effectiveness of interventions. Most of the portals, particularly those based in the United States, were focused on the effectiveness of particular branded programmes. This has some logic as they are available interventions and may have undergone a period of research and development focused on particular social policy issues. On the other hand the programme approach provides rather narrow research questions. The programmes can be compared for relative effectiveness but if the differences are small then how can users choose between them? It might be more productive to follow the approach of some portals (e.g. EEF, NICE, WW Growth) to assess the evidence base of broader definitions of approaches. It might also be helpful to have issue driven questions that include logic models and process data so that there can be greater clarity about causal processes and how certain interventions may be relatively more or

less effective in different contexts and how interventions can be adapted for such contexts (Kneale et al 2018). The challenge in this approach is to code branded programmes in order to assign them to one of the more broadly defined approaches. A single programme may combine multiple components – though with sufficient sample size exploring additive or interaction effects of such combinations should be possible.

It is also worth noting that the portals taking a 'programme perspective' were also the portals most likely to take a 'one or more good studies' approach to making statements about an evidence base. This is not surprising in the sense that the portals were set up to assess whether there was evidence in support of a programme rather than a comparative analysis of relative efficacy of different approaches to a problem (see Evidence Base General method and Recommendation 3 below).

Recommendation 2

Consider using broad rather than narrow evidence base questions

(i) Consider using broad research questions that include evidence on effectiveness, cost effectiveness, and theory and data about the process by which impact is achieved in different contexts. This could involve: (a) an analysis of the problem or issue; (b) consideration of a theory of change as to why certain types of interventions might help – this might require a review of research on causal processes; (c) a review of the evidence of the cost effectiveness of interventions to achieve the desired benefit in different contexts.

(ii) Take care about the fitness for purpose of narrow 'black box' questions of efficacy of individual interventions without consideration of the broader context of similar interventions.

3. Evidence base general method

The portals varied in their approaches to appraising evidence bases. Some portals used pre-existing systematic reviews that used explicit rigorous methods to review the combined findings of already existing studies (see Section 2.1). Some undertook new systematic or non-systematic reviews and some made judgements on the basis of one or more rigorous evaluations (Gough et al 2017).

The strongest evidence (and thus the highest evidence standard) arises from systematic reviews as these use a transparent research process to identify what is or is not known about an issue from previous research (the evidence base at the time the review was undertaken).

This is not to suggest that all systematic reviews are fit for purpose. Even if they are excellent in terms of method they may not be framed in a way that helps address the decisions that are being made. This is particularly problematic if the synthesis report is not clear about the way the research question was framed (and the resultant way that the synthesis was undertaken) resulting in the findings being misunderstood and misapplied by decision makers. What is required is a synthesis that is rigorous, explicitly reported and relevant to the decisions being made. As synthesis of research may be of generalised global knowledge then there may always need to be interpretation of the evidence for local contexts.

If there are pre-existing systematic reviews then there is efficiency and less research waste in using those. Care, however, is needed to ensure that the reviews sufficiently fit the technical standards and the relevance of the portal's research question. An alternative is to commission new reviews, but these need to be systematic (in being rigorous and transparent in method) just as is required of primary research and so meet the evidence standards for reviews (Gough 2016, Liabo et al 2017).

The reliance on one or two good quality rigorous studies seems full of dangers. It takes no account of other studies showing either no effect or harmful effect. It is therefore weaker than even a vote counting approach that counts up the number of studies reporting positive, negative, or no effects, which has well known shortcomings.

Studies of effectiveness are based on probability estimates and so individual studies may, however well executed, be prone to random error and thus misrepresent the underlying reality. An examination of a forest plot from in statistical meta-analysis clearly shows the way that individual studies vary in their results. Relying on one or two of the studies that happen to show a positive finding without reference to other relevant studies may be very misleading. An aggregation of the data in a statistical synthesis provides an indication of the total evidence base rather than studies selected on the basis of their results. Another example of this is the accusation that some providers of an intervention undertake many outcome studies on their products until the desired result is obtained. Even without this conflict of interest, outcome reporting bias in favour of positive outcomes is well documented for both single studies and non-systematic literature reviews.

Even where there are very few studies on a topic, it can still be useful to use systematic review methods to map and synthesise what studies have been undertaken. If we are rigorous in specifying what has been studied and what we do (or do not) know and how we know it, then we are in a better position to plan what more that we need to know (and how we could know it). Reviews therefore can help us be evidence informed about planning new research.

Recommendation 3

Use explicit rigorous methods of evidence synthesis to make claims about the existing evidence base

(i) Prioritise the use of systematic reviews to make evidence claims about an evidence base in order of:

- Pre-existing review if rigorous and relevant.
- Review of pre-existing reviews if rigorous and relevant (and overlap and infilling possible).
- New systematic review(s).

(ii) Avoid non-systematic reviews and 'one or two good studies' methods.

(iii) Specify and justify the review methods used and follow all of the expected stages, processes and transparency expected of systematic review including:

- The type of review being undertaken.
- The nature of the review questions being asked and any underlying perspectives and assumptions.
- The study methods inclusion criteria. Narrow questions of efficacy will require rigorous studies of impact which are likely to be experimental studies controlling for the effects of extraneous variables (counterfactuals). Non experimental studies such as natural experiments may also provide useful data where sample sizes are very large and the effects of variables can be assessed or to provide background prevalence data.
- Other inclusion criteria such as: specification of the topic and geographical, historical and language limits.

- The method of coding and analysing (synthesising) the data from included studies. Consider allowing the evidence to be filtered by types of intervention, outcome, and context though recognising that the strength of evidence will be less for such sub-group analysis.

5. Evidence bases: data and evidence standards

The portals used a variety of approaches to considering the appraisal of the whole evidence base. All of these can be considered as useful data to inform decision making.

Recommendation 4

Specify and justify the different evidence standards for making different claims about the existing evidence base including impact, strength, extent and consistency of evidence, process and contexts and costs.

Consider the specification and justification of the evidence standards in terms of:

- Impact both beneficial and harmful (the effect of the intervention).
- Strength, extent and consistency of evidence (how robust and trustworthy).
- Nature of outcome variables measured and the follow up time period (how was the effectiveness of the intervention measured?).
- Process data and logic models to understand cause and thus interpretation of causal impact.

- Variation between studies in the context in which they were undertaken (and to which the research findings might therefore apply).
- Cost and cost-benefit data (the resources and how their relative cost compared to the beneficial effect of the intervention as stated in the evidence base).

6. Included studies: inclusion criteria and evidence standards

All of the portals specified the types of studies that would help them address the question of effectiveness. Most of them did this in quite a broad way. This has the advantage of not mistakenly missing an important study though may result in greater efforts in screening a large number of potential includes from the search. More important methodologically is the appraisal of the evidence standards for included studies.

Recommendation 5

Specify and justify evidence standards for included studies

Specify and justify the evidence standards (and/or inclusion criteria) for included studies so that they are: (a) fit for purpose for the review question being addressed; (b) consistent with the evidence standards applied to the evidence claims of the whole review (as in Recommendations 1 and 5).

7. Process versus criteria for evidence standards

The ways in which the portals explained their methods and evidence standards varied considerably in both approach and detail. EIS and WW Gowth, for example, provide great detail about the criteria for appraising primary studies but less on how to synthesise and appraise the evidence base of all relevant studies. EIF provides a scale of the criteria for appraising one or more studies to make a statement about the evidence base but provides relatively little detail compared to EIS on the methodological issues involved. NICE provides great detail on the methods for synthesising studies in an evidence base and proposes, but does not require, specific criteria for appraising the final evidence on effectiveness.

So what is the best way of presenting standards? Is it a one page list of criteria? Or is it a lengthy detailed manual specifying both the method and process of how rigorous methods can be achieved? Maybe the best solution is a balance between specification of method, process and criteria, as well as the balance between reviews of an evidence base and of the studies that are part of that evidence base.

Recommendation 6

Specify methods as well as criteria for achieving evidence standards

Consider providing transparency on:

- Methodological steps in creating evidence of a certain standard.
- Procedures for the methods to be undertaken (including internal and external quality assurance processes).
- Statements and scales specifying evidence standards to be reached.

8. Guidance and recommendations for decision makers

Guidance and recommendations provide interpretations of what the evidence means for policy and practice and individual decisions. As the purpose of portals is to engage research use with research production then evidence informed guidance may be the most useful product for busy decision-makers. It may be more efficient for committees representing the main stakeholders to engage with the details of the research and interpreting its implications rather than every decision maker doing this individually. This may also increase the focus on the processes by which interventions may have their effect and how these may be effected by local contexts in which policy and practice is applied.

A further benefit of guidance development processes is that they can enable user perspectives to influence research questions (as in, for example, starting with challenges that decision makers are trying to address rather than asking about the effectiveness of intervention programmes). The process of creating such guidance has become more explicit with the processes of the GRADE framework and the guidance products of EEF, IES and NICE. Guideline producers can also draw on the work of guideline organisations such as the Guidelines International Network.

Guidance and recommendations can therefore be a useful or the most useful product of an evidence portal. This then raises the issue of the evidence standards for such guidance. There is, for example, much more that could be done to be systematic and transparent about the involvement of different stakeholders, how they are involved, social values of evidence appraisal, and the types of non-evidence base information that is used in different ways in moving to making recommendations.

Recommendation 7

Develop methods and standards for policy and practice guidance informed by the evidence base

Consider the potential use and benefits of evidence informed guidance and how it is created through specifying:

- The role that guidance could play within the theory of change of the portal.
- The engagement of different stakeholder groups.
- The types of information considered and the evidence standards for that information and for how it is combined with all the other types of evidence considered (including the effectiveness research evidence base).
- The social values underlying guidance.

Implications for the Centre for Homelessness Impact

All of the above recommendations will inform the further development of CHI's own portal - the Intervention Tool.

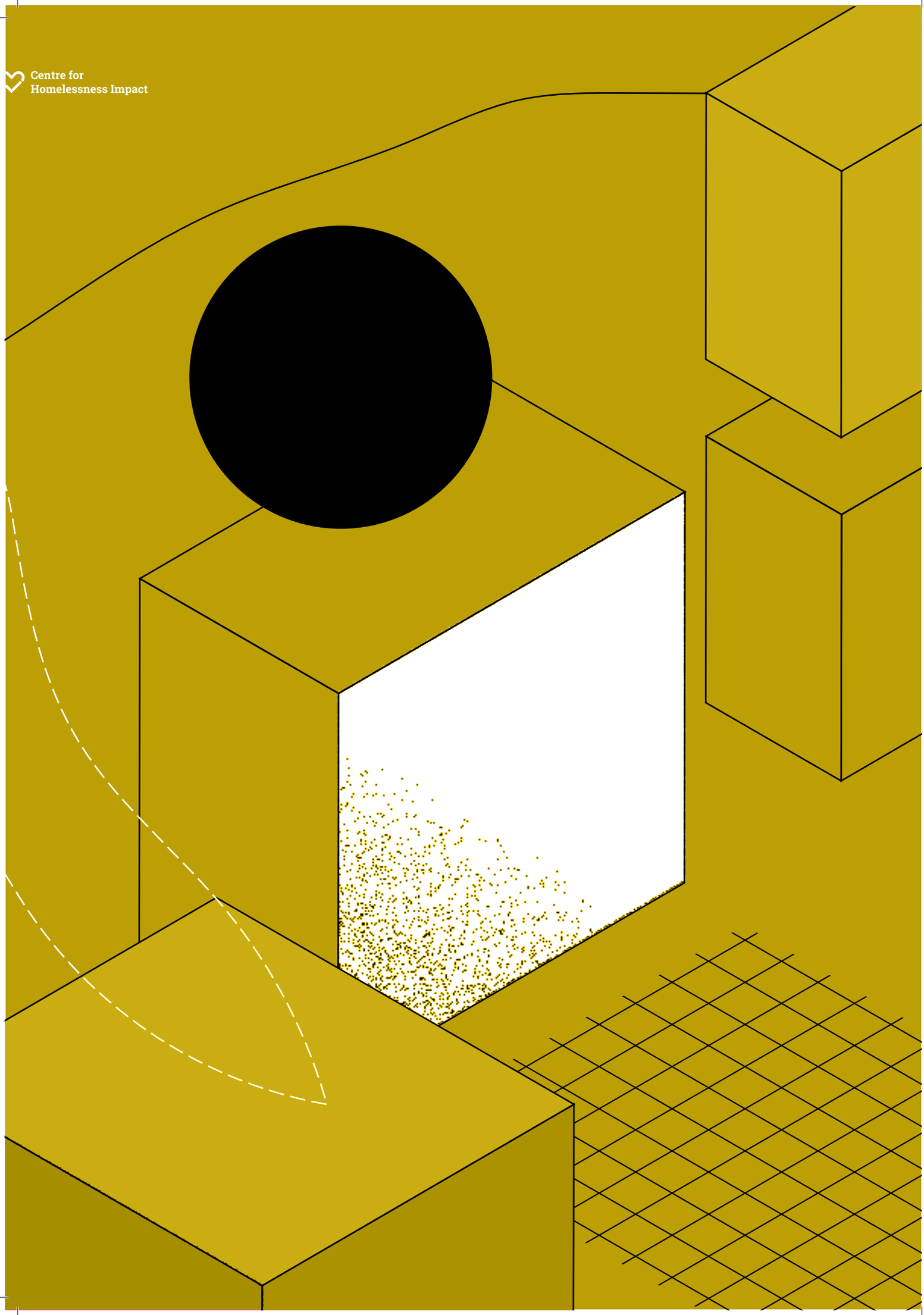
The tool was modelled after EEF's as it appeared to be more successful than others and their end users seemed to prefer it when presented with different options. The beta version of the tool includes some important modifications in response to user need (e.g. information about where the evidence comes from and impact on a variety of outcomes) and soon a second version will be released which includes information about implementation issues.

Comprehensive user testing will also enable CHI to embed the tool in day-to-day business processes and understand how it works and how it could be further developed to meet the needs of the end users of the research. From what we know about how these portals this is vital to their success.

In terms of standards of evidence CHI is committed to draw on systematic review evidence. CHI has used the evidence map produced with Campbell to identify priority review topics and over time all of the portal will be populated using evidence from systematic reviews.

Doing so will take time as systematic reviews take a lot longer to produce than other types of reviews and they are also a lot more expensive. To ensure the needs of the end users of the evidence are met in the meantime, the Centre in partnership with Campbell populated the content of the tool with summaries from the Evidence and Gap Maps.

Centre for
Homelessness Impact



References

Department for International Development (DfID) (2014). Assessing the Strength of Evidence. How to Note. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf

Gough DA (2016) Evidence Standards: A Dimensions of Difference Framework for Justifiable Evidence Claim (with special reference to systematic reviews). London: EPPI-Centre, University College London.

Gough D, Maidment C, Sharples J (2018). UK What Works Centres: Aims, methods and contexts. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London. ISBN: 978-1-911605-03-4

Gough D, Oliver S, Thomas J. (2017), Introduction to systematic reviews. 2nd Edition. London: Sage.

Kneale D, Thomas J, Bangpan M, Waddington H, Gough D(2018) Conceptualising causal pathways in systematic reviews of international development interventions through adopting a causal chain analysis approach, Journal of Development Effectiveness, DOI: https://doi.org/10.1080/19439342.2018.1530278"10.1080/19439342.2018.1530278

Langer L, Tripney J, Gough D (2016). The Science of Using Science: Researching the Use of Research Evidence in Decision-Making. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

Liabo K, Gough D, Harden A (2017) Developing justifiable evidence claims. In: Gough D, Oliver S, Thomas J. (eds) Introduction to systematic reviews. 2nd Edition. London: Sage.

Appendix 1

Platform Descriptions

- California Evidence-Based Clearinghouse for Child Welfare (CEBC)
- Clearinghouse for Labor Evaluation and Research (CLEAR)
- Conservation Evidence
- Early Intervention Foundation (EIF) Guidebook
- Education Endowment Foundation (EEF) Toolkit
- European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) Practice Portal Evidence Database
- European Platform for Investing in Children (EPIC): Evidence Based Practices
- Evidence Based Teen Pregnancy Programs (EBTT)
- Institute of Educational Sciences (IES) What Works Clearing House
- National Institute for Health and Care Excellence (NICE)
- What Works Centre for Crime Reduction: Crime Reduction Toolkit
- What Works Centre for Local Economic Growth Toolkits
- What Works in Reentry Clearinghouse
- What Works Wellbeing Evidence Comparison Tool

California Evidence-Based Clearinghouse for Child Welfare (CEBC)

Topics

The mission of the California Evidence-Based Clearinghouse for Child Welfare is to advance the effective implementation of evidence-based practices for children and families involved with the child welfare system. Topic groups include:

- Anger management, domestic violence, substance abuse (with 5 topics)
- Behaviours management, including parent training (with 7 sub-topics)
- Child welfare services, including placement and reunification (with 13 topics)
- Engagement and Parent Partnering Programs (with 4 topics)
- Mental health (with 9 topics)
- Prevention and early intervention (with 6 topics)
- Supportive services for youth in the child welfare system (with 6 topics)

Metrics

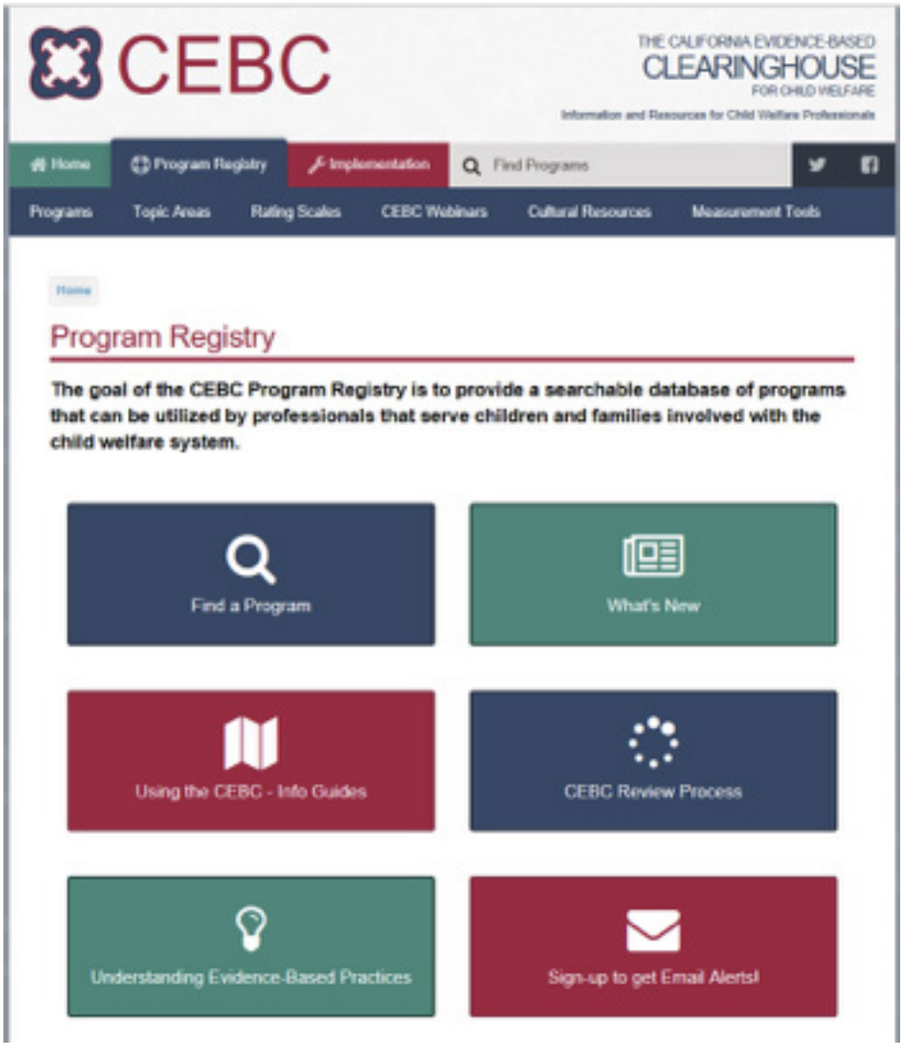
For each intervention there is information on:

- Target population
- Evidence rating
- Child Welfare System Relevance Level with 3 levels of: High, Medium and Low as being as designed, or is commonly used, to meet the needs of children, youth, young adults, and/or families:
- Receiving child welfare services. (High)
- Similar to child welfare populations and likely include current and former child welfare services recipients. (Medium)
- With little or no apparent similarity to the child welfare services population. (Low)

- Programme information: About This Program; Brief Description; Essential Components; Child/Adolescent Services; Parent/ Caregiver Services; Recommended Parameters; Delivery Settings; Homework; Languages; Resources Needed to Run Program; Minimum Provider Qualifications; Education and Training Resources; Implementation Information; Relevant Published, Peer-Reviewed Research; References; Contact information

There are not quality or extent of evidence ratings for individual primary studies.

Screeenshot
http://www.cebc4cw.org/home/



Website pages and searching

Clearinghouse website: http://www.cebc4cw.org/
Topic areas: http://www.cebc4cw.org/search/topic-areas/
Review and rating process: http://www.cebc4cw.org/home/how-are-programs-on-the-cebc-reviewed/
Scientific Rating Scale: http://www.cebc4cw.org/ratings/scientific-rating-scale/

The Program Registry page allows advanced searching by:

- Keyword
- Scientific Rating
- Child Welfare System Relevance Level
- Child Welfare Outcomes
- Topic Areas
- Age of Child

Clearinghouse for Labor Evaluation and Research (CLEAR)

Topics

CLEAR reviews studies in a variety of labor-related topic areas that are determined in collaboration with the DOL Chief Evaluation Office,DOL agencies, CLEAR contractors and CLEAR technical working groups of advisors. Topics covered to-date include:

- Apprenticeship and On-the-Job Training: studies that look at work-based training programs' effectiveness on participants' employment and earnings outcomes.
- Behavioral Finance: Retirement: studies that examine interventions grounded in insights from behavioral economics that may encourage employees to save more for retirement.
- Behavioral Insights: studies of interventions that apply behavioral science insights to labor-related contexts.
- Career Academies: studies that examine a high school

intervention model first launched in 1969 that, as of 2013, served approximately one million students in 7,000 schools nationwide.

- Child Labor: studies that examine interventions intended to reduce or prevent child labor for children under 18.
- Community College: studies that look at community college interventions that are intended to improve academic persistence, degree/certificate completion, and labor market outcomes of community college students.
- Disability Employment Policy: studies that examine programs that seek to improve labor market outcomes for youth and working-age adults with disabilities.
- Employer Compliance: studies that examine rules, policies, and enforcement activities aimed at preventing discrimination by employers.
- Entrepreneurship and Self-Employment: studies that look at interventions that promote individuals' entrepreneurship and self-employment.
- Job Search Assistance: studies of job search assistance interventions, which aim to improve participants' employment and earnings outcomes.
- Low-Income Adults: studies that examine interventions that serve low-income adults with the goal of improving their employment and earnings outcomes.
- Older Workers: studies of employment and training programs and broad federal or state interventions that support and/or improve the employment prospects and financial security for workers age 40 and older.
- Opportunities for Youth: studies of interventions designed to provide opportunities for youth aged 14-24 who have not recently been in school or the labor force to improve their labor market outcomes.
- OSHA Enforcement: studies that examine the effectiveness of Occupational Health and Safety Administration (OSHA) enforcement activities.
- Reemployment: studies that examine interventions designed to promote faster reemployment of unemployment insurance claimants.
- Reentry: studies of employment and training programs that encourage basic skills development, educational attainment,

employment, employment retention, and career advancement for individuals returning from incarceration.
- Veterans: studies of employment and training programs that encourage basic skills development, educational attainment, completion of training programs and/or acquisition of certificates or credentials, employment, employment retention, and career advancement for veterans.
- Women in Science, Technology, Engineering, & Math (STEM): studies that examine interventions designed to foster success among women in STEM fields.

Metrics

What's In a Profile:

CLEAR highlights key features of all the relevant research identified for a given topic area, including causal analyses as well as descriptive analyses and implementation studies in a study profile.

For descriptive analyses and implementation studies, the profiles contain a concise summary of study "highlights," including the:

1. Research question
2. Intervention and setting
3. Data and methods
4. Findings

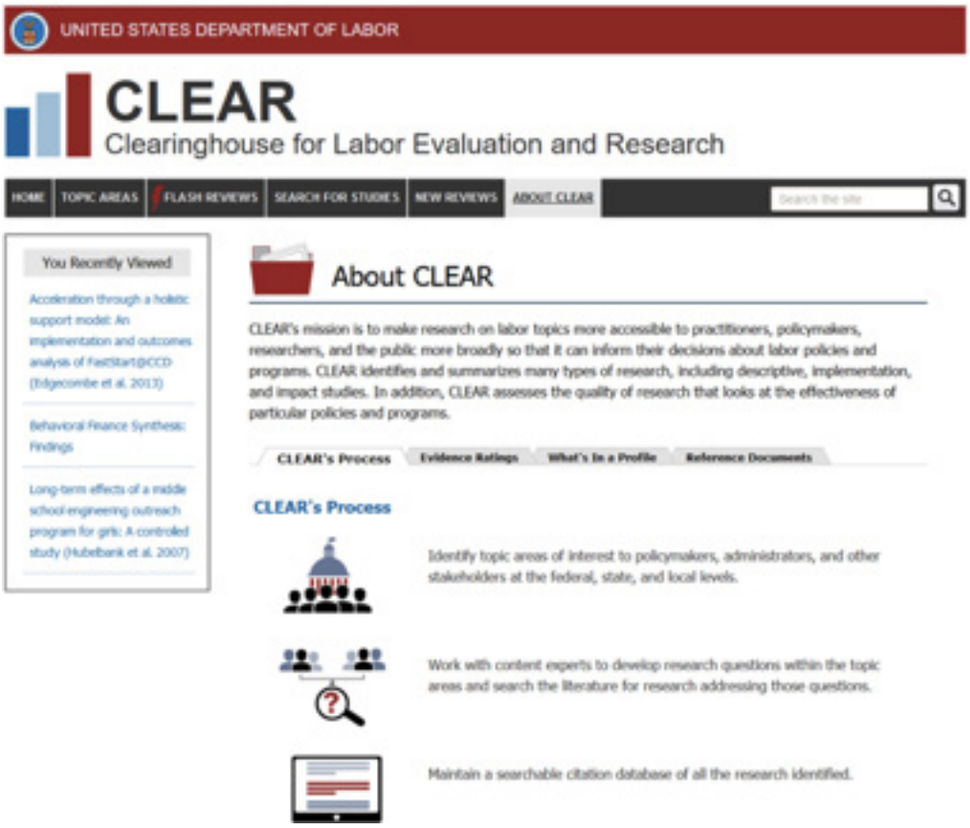
For causal studies that estimate impact, CLEAR provides ratings of study quality and produces more in-depth research profiles. These types of studies, when they are of sufficient quality, can answer questions about if an intervention "works" for practitioners and policymakers. For each causal study reviewed, these profiles present information on:

- Features of the intervention or programme
- Features of the study (data and methods)
- Findings
- Considerations for interpreting the findings
- Causal evidence rating and an explanation for why the research received the rating

Where there is a synthesis of the literature, the portal presents:

- Key findings across the evidence base
- Gaps in the literature
- List of studies

Screeenshot
https://clear.dol.gov/about



Website pages and searching
Clearinghouse website: https://clear.dol.gov/
Policies and Procedures Version 3.1: https://clear.dol.gov/reference-documents/policies-and-procedures-version-31
Causal Evidence Guidelines, Version 2.1: https://clear.dol.gov/reference-documents/causal-evidence-guidelines-version-21

Study Review Guide for Causal Reviews: https://clear.dol.gov/reference-documents/study-review-guide-causal-reviews
Quantitative Descriptive Guidelines: https://clear.dol.gov/reference-documents/quantitative-descriptive-guidelines
Implementation Study Guidelines can be found here: https://clear.dol.gov/reference-documents/clearinghouse-labor-evaluation-and-research-clear-guidelines-reviewing-0

The search for studies page allows searching by keywords, topic area, study type and review type.

The site discusses topic selection and provides detail on process and a causal study review coding guide

Conservation Evidence

Topics

Conservation Evidence is an authoritative information resource summarising evidence from the scientific literature about the effects of conservation interventions, such as methods of habitat or species management. Main categories to date are:

Bird Conservation; Primate Conservation; Control of Freshwater Invasive Species; Amphibian Conservation; Peatland Conservation; Forest Conservation; Farmland Conservation; Shrubland and Heathland Conservation; Management of Captive Animals; Bat Conservation; Mediterranean Farmland; Bee Conservation; Soil Fertility; Sustainable Aquaculture; Natural Pest Control.

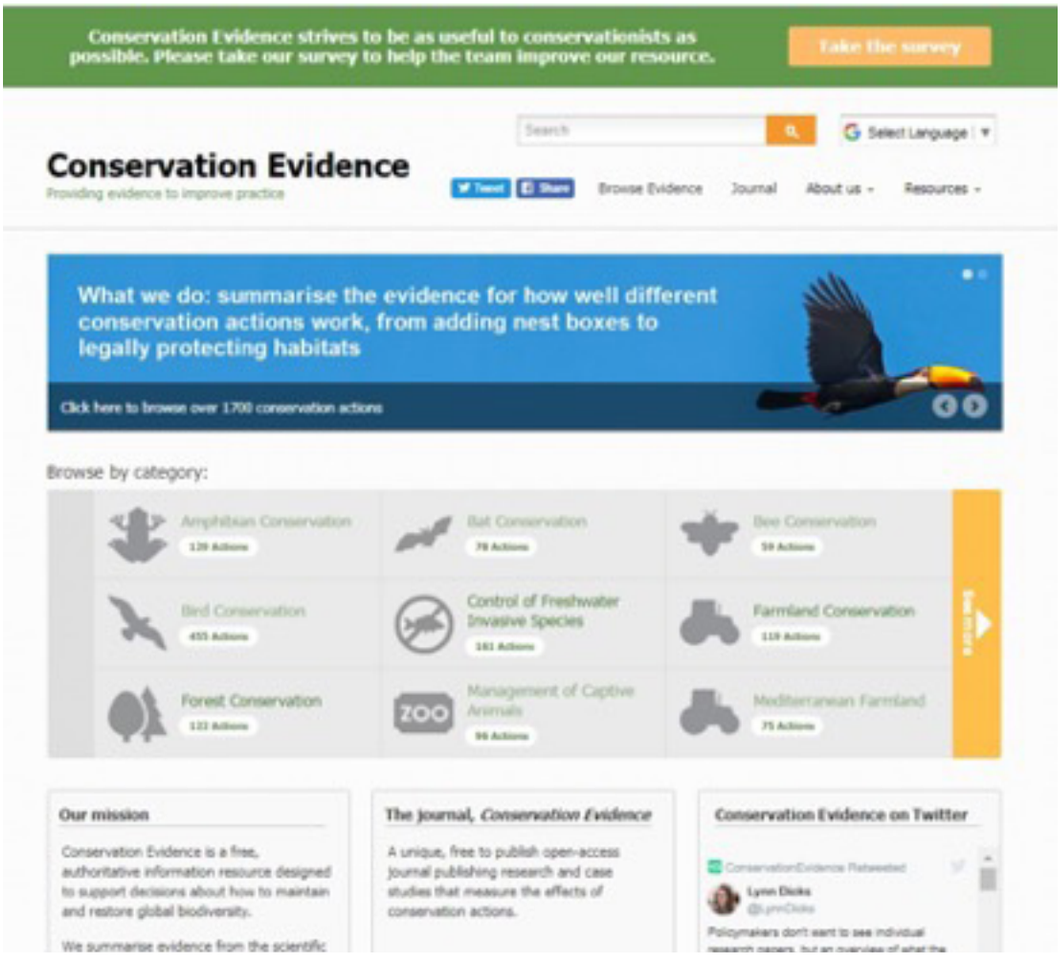
Metrics

For each action there is information on:

- The studies found testing the action
- Their findings
- Key messages
- Category of effectiveness, based on % scores for Effectiveness, Certainty and Harms by a group of experts

Screeenshot

https://www.conservationevidence.com/



Website pages and searching

Homepage: https://www.conservationevidence.com/
Methods: https://www.conservationevidence.com/site/page?view=methods
Actions: https://www.conservationevidence.com/data/index
Individual studies: https://www.conservationevidence.com/data/studies?path=data%2Findex
What is What Works in Conservation?: https://www.conservationevidence.com/content/page/79

Evidence on action example: https://www.conservationevidence.com/actions/398

Search by:

- Keywords
- Country
- Category (currently): Bird Conservation; Primate Conservation; Control of Freshwater Invasive Species; Amphibian Conservation; Peatland Conservation; Forest Conservation; Farmland Conservation; Shrubland and Heathland Conservation; Management of Captive Animals; Bat Conservation; Mediterranean Farmland; Bee Conservation; Soil Fertility; Sustainable Aquaculture; Natural Pest Control
- Habitat: Artificial Habitats; Forest & Woodland; Wetlands; Shrubland; Grassland; Rivers, Streams, Creeks; Coasta; Marine; Savanna; Rocky Habitats & Caves; Desert; Other
- Threat: Agriculture & aquaculture; Invasive & other problematic species & genes; Biological resource us; Natural system modifications; Pollution; Residential & commercial development; transportation & service corridors; Energy production & mining; Human intrusions & disturbance; Climate change & severe weather; Geological events; Other
- Action type: Land/water management; Species management; Livelihood, economic & other incentives; Education and awareness; Law & policy; Land/water protection; Other

Early Intervention Foundation (EIF) Guidebook

Topics

TThe EIF Guidebook provides information about 82 early intervention programmes that have been evaluated and 45 of which have shown to improve outcomes for children and young people (to be effective on Levels 3 and 4 of EIF evidence rating). EIF has rated the strength of evidence for a programme's impact and its relative costs. The Guidebook also provides information about the specific outcomes a programme has been shown to improve, how the programme works, how it is delivered, and the conditions or resources that can make a programme more likely to be effective.

Metrics

For each intervention the Guidebook presents information on:

- Evidence rating (strength of evidence of effectiveness from the evidence base)
- Cost rating
- Child outcomes
- Key programme characteristics: Age group it is for, Delivery model, Delivery setting
- Classification (Targeting): Universal, Targeted Selected, Targeted Indicated
- About the programme: how does it work / theory of change

Screeenshot
https://guidebook.eif.org.uk/eif-evidence-standards/

Website pages and searching

EIF Guidebook: http://guidebook.eif.org.uk/
How to read the Guidebook: http://guidebook.eif.org.uk/guidebook-help/how-to-read-the-guidebook#what-is-the-evidence-rating

The home page allows searching by evidence rating, cost rating, if implemented in the UK, child outcomes, age group. The report page provides links to:

Foundation for Life webpage with links to report and basis of Guidebook cost ratings: http://www.eif.org.uk/publication/foundations-for-life-what-works-to-support-parent-child-interaction-in-the-early-years/

Education Endowment Foundation (EEF) Toolkit

Topics

The Education Endowment Foundation Teacher and Learning toolkit provides an accessible summary of the international research evidence on teaching 5 to 18 year olds. It shows the evidence for 34 intervention categories for school education such as 'Setting or streaming', 'School uniform' and 'Arts Participation'. It is planned to expand the number of categories over time.

Metrics

For each intervention the toolkit presents information on:

- Cost: shown by a £ symbol with a range of between £1 and £5
- Evidence strength: shown by a lock, with a range of between 1 and 5 locks
- Impact (months): months additional learning outcomes from exposure to the intervention

Screeenshot
https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit

Website pages and searching

EEF homepage: https://educationendowmentfoundation.org.uk/
Evidence summaries: https://educationendowmentfoundation.org.uk/evidence-summaries/
Example of a technical appendix; https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/arts-participation/technical-appendix
Using our toolkits: https://educationendowmentfoundation.org.uk/evidence-summaries/about-the-toolkits/
Technical appendices: https://educationendowmentfoundation.org.uk/public/files/Toolkit/Toolkit_Manual_2018.pdf
The Teaching and Learning Toolkit: what are the risks and how do we address them?: https://educationendowmentfoundation.org.uk/news/the-teaching-and-learning-toolkit-what-are-the-risks-and-how-do-we-address/

The evidence may be sorted by each metric (with alphabetical sort of intervention category as the secondary sort field).

The homepage also provides a filter to filter programmes by the three metrics (left hand panel In Figure 1).

Clicking on an intervention category leads to a page which provides:

- A brief description of the intervention
- A brief discussion impact including heterogeneity (where impact is stronger or weaker)
- The quality of the evidence
- Data on costs, noting variation in cost estimates
- Other factors to consider in selecting the intervention
- Technical appendix

The page also provides links to:

1. A summary of the review
2. The full evidence review
3. Related projects funded by EE
4. Further reading (links to other relevant resources)

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA)

EMCDDA Best Practice Portal Evidence Database

Topics

This database gives you access to the latest evidence on drug-related interventions. All of the many topics covered by the website (not just the evidence database) are listed alphabetically at: http://www.emcdda.europa.eu/topics_en.

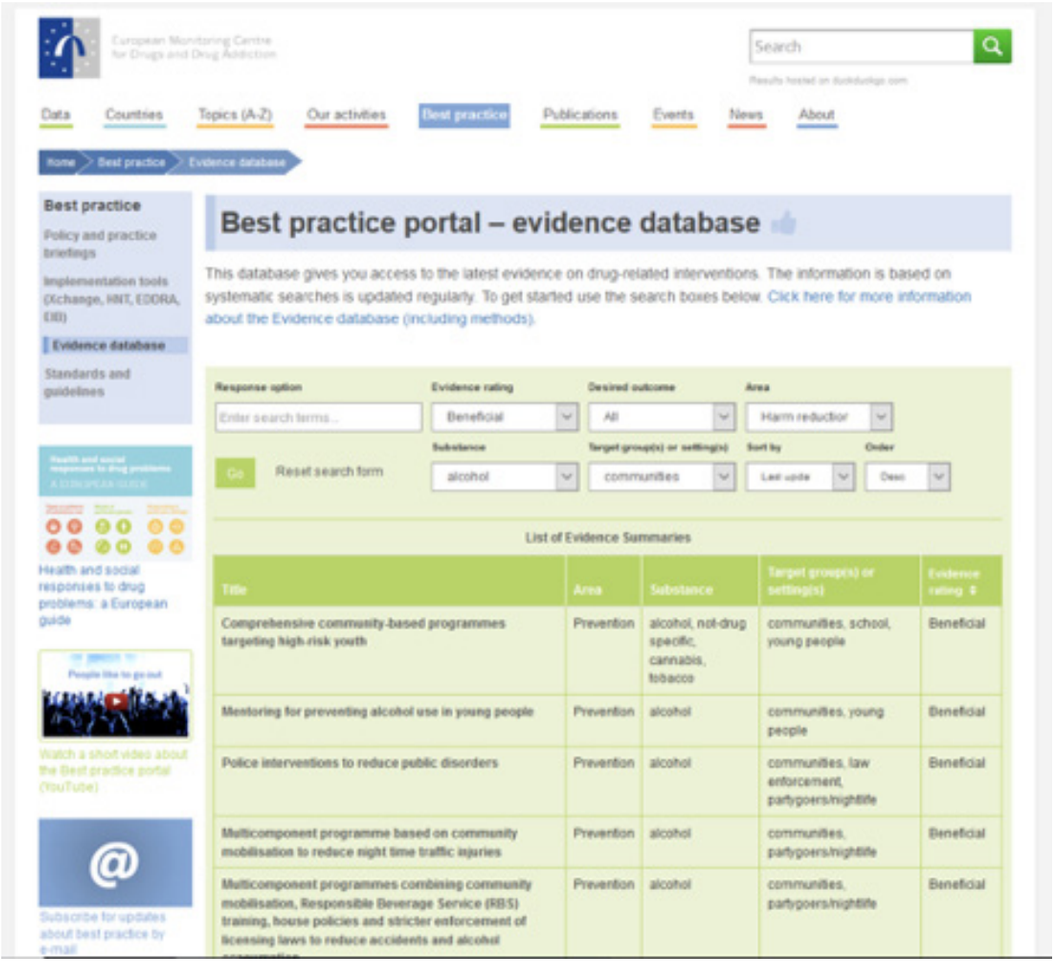
Metrics

For each intervention the toolkit presents the overall evidence rating, lists the reviews on which this based including a summary of their findings.

- Cost: shown by a pound symbol with a range of between 1 and 5 pounds
- Evidence strength: shown by a lock, with a range of between 1 and 5 locks
- Impact (months): months additional learning outcomes from exposure to the intervention

Screeenshot

http://www.emcdda.europa.eu/best-practice/evidence-summaries_en?response_option=&evidence_rating=1181&outcome=All&evidence_summary_area=All&substance=1027&target=1070&sort_by=field_pub_date_value&sort_order=DESC



Website pages and searching

Homepage: http://www.emcdda.europa.eu/best-practice/evidence-summaries_en
Evidence Ratings: http://www.emcdda.europa.eu/best-practice/evidence/about
Best Practice Portal: http://www.emcdda.europa.eu/best-practice
Example of an evidence summary: http://www.emcdda.europa.eu/best-practice/evidence-summaries/police-interventions-reduce-public-disorders_en

The Best practice portal evidence database page provides searching for evidence by:

- Open search
- Evidence rating: Beneficial; Likely to be beneficial; Evidence of ineffectiveness; Trade-off of benefits and harms; Unknown effectiveness
- Desired outcome: multiple
- Area: Emerging topics; Harm reduction; Other topics; Prevention; Social reintegration; Treatment
- Substance: multiple
- Target group or setting: multiple

European Platform for Investing in Children (EPIC): Evidence Based Practices

Topics

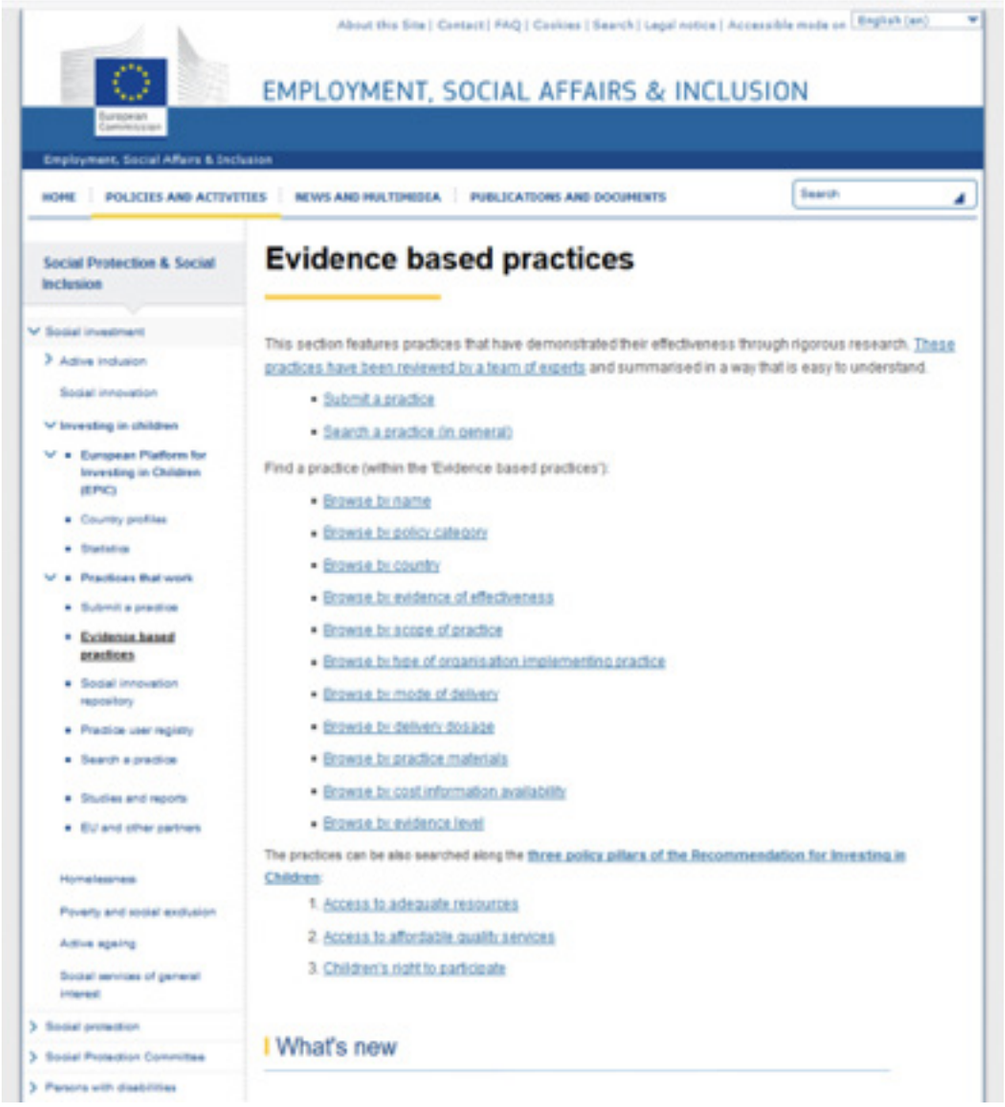
The European Platform for Investing in Children (EPIC) is an evidence-based online platform that provides information about policies that can help children and their families face the challenges that exist in the current economic climate in Europe. The platform has evidence pages on 62 practices.

Metrics

For each intervention the toolkit presents information on:

- Project overview
- Recommendation pillars
- Countries that have implemented practice
- Age groups
- Target groups
- Years in operation
- Type of organisation implementing practice
- Transferability
- Evidence of effectiveness
- Issues to consider
- Years in operation
- Available resources
- Evaluation details
- Enduring impact
- Bibliography
- Contact information

Screeenshot
http://ec.europa.eu/social/main.jsp?catId=1251&langId=en



Website pages and searching

Homepage: http://ec.europa.eu/social/main.jsp?catId=1251&langId=en
The Evidence based practices page: http://ec.europa.eu/social/main.jsp?catId=1251&langId=en

EPIC Evidence-Based Practices - Review criteria and process: http://ec.europa.eu/social/main.jsp?catId=1246&intPageId=4286&langId=en
Example of evidence page on an intervention: http://ec.europa.eu/social/main.jsp?catId=1251&langId=en&reviewId=188

The Evidence based practices page allows searching by:

- Name
- Policy category
- Country
- Evidence of effectiveness
- Scope of practice
- Type of organisation implementing practice
- Mode of delivery
- Delivery dosage
- Practice materials
- Cost information availability
- Evidence level

The practices can be also searched along the three policy pillars of the Recommendation for Investing in Children:

1. Access to adequate resources
2. Access to affordable quality services
3. Children's right to participate

Evidence Based Teen Pregnancy Programs

Topics

44 evidence-based teen pregnancy prevention (TPP) programs that have been shown, in at least one program evaluation, to have a positive impact on preventing teen pregnancies, sexually transmitted infections, or sexual risk behaviours.

Metrics

For each intervention there is information on:

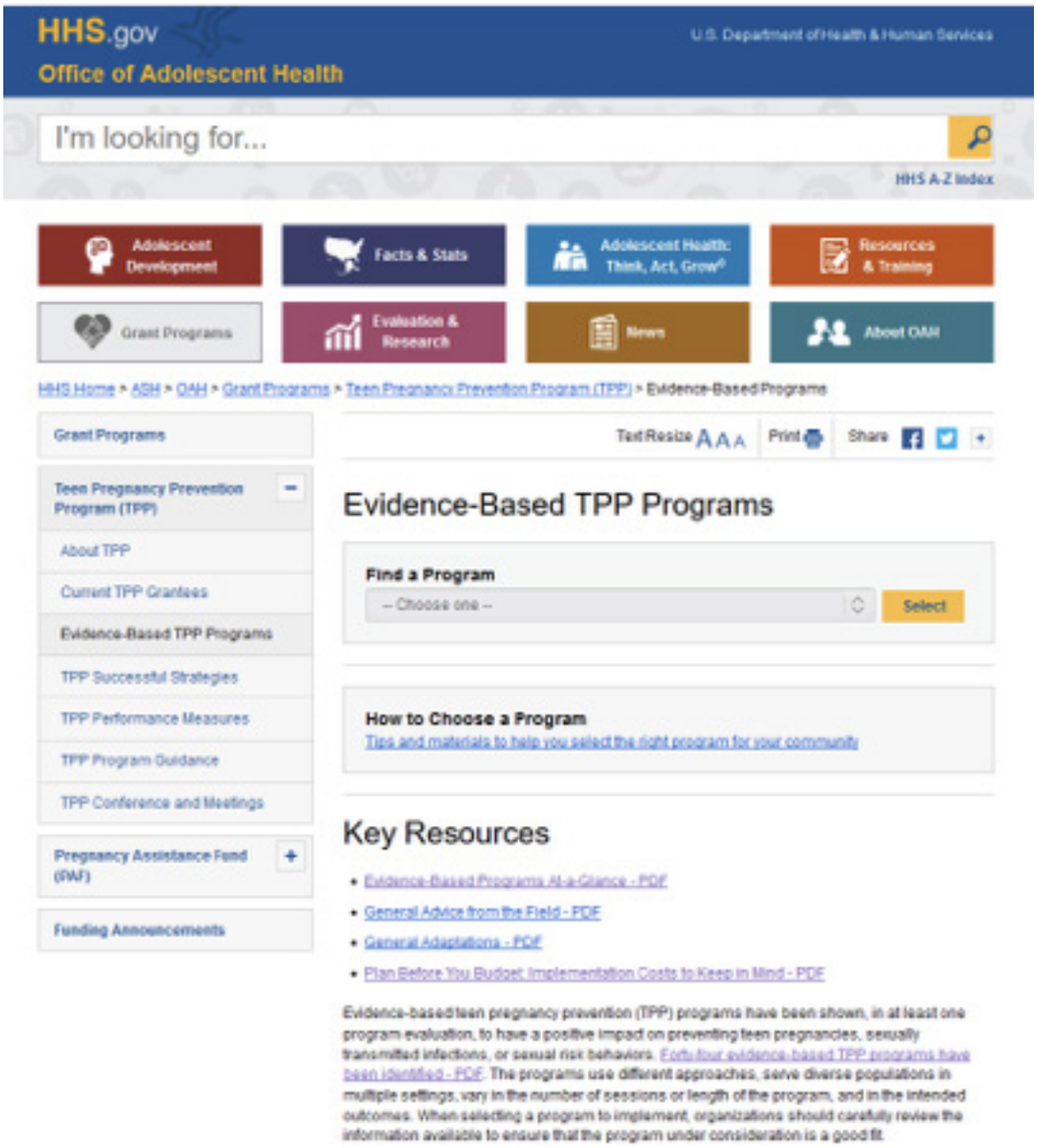
Progam Information:

- Program Overview: Developer(s); Program Summary; Target Population; Program Setting; Contact and Availability Information.
- Program Components: Program Objectives; Program Content; Program Methods.
- Implementation Requirements and Guidance: Program Structure and Timeline; Staffing; Program Materials and Resources; Additional Needs for Implementation; Fidelity; Training and Staff Support; Allowable Adaptations.
- Implementation Readiness Assessment.

Research Evidence:

- Reviewed Studies
- Study Characteristics:
- Study findings: Evidence by outcome domain and study: Detailed findings

Screeenshot https://www.hhs.gov/ash/oah/grant-programs/teen-pregnancy-prevention-program-tpp/evidence-based-programs/index.html



Website pages and searching

Website: https://www.hhs.gov/ash/oah/grant-programs/teen-pregnancy-prevention-program-tpp/evidence-based-programs/index.html
Teen Pregnancy Prevention Evidence Review: Review process: https://tppevidencereview.aspe.hhs.gov/ReviewProtocol.aspx
Review Protocol 5.0: https://tppevidencereview.aspe.hhs.gov/pdfs/TPPER_Review%20Protocol_v5.pdf
Teen Pregnancy Prevention Evidence Review: Families Talking Together: https://tppevidencereview.aspe.hhs.gov/document.aspx?rid=3&sid=53&mid=5

The home page allows searching by programme name. This provides link to a programme page with information on the components of the programme, target populations and advice from the field on programme components and stakeholder and parent buy-in, There is also a link 9at the very bottom of the page) to the research evidence on the programme - for example: Read the research about the PHAT-AO program at ASPE's website.

Institute of Education Sciences (IES): What Works Clearinghouse

Topics

Intervention reports on nine topic areas and 3 age groups with total of 581 intervention reports (7th November 2018): interventions are very specific, usually specific branded programnes.

Metrics

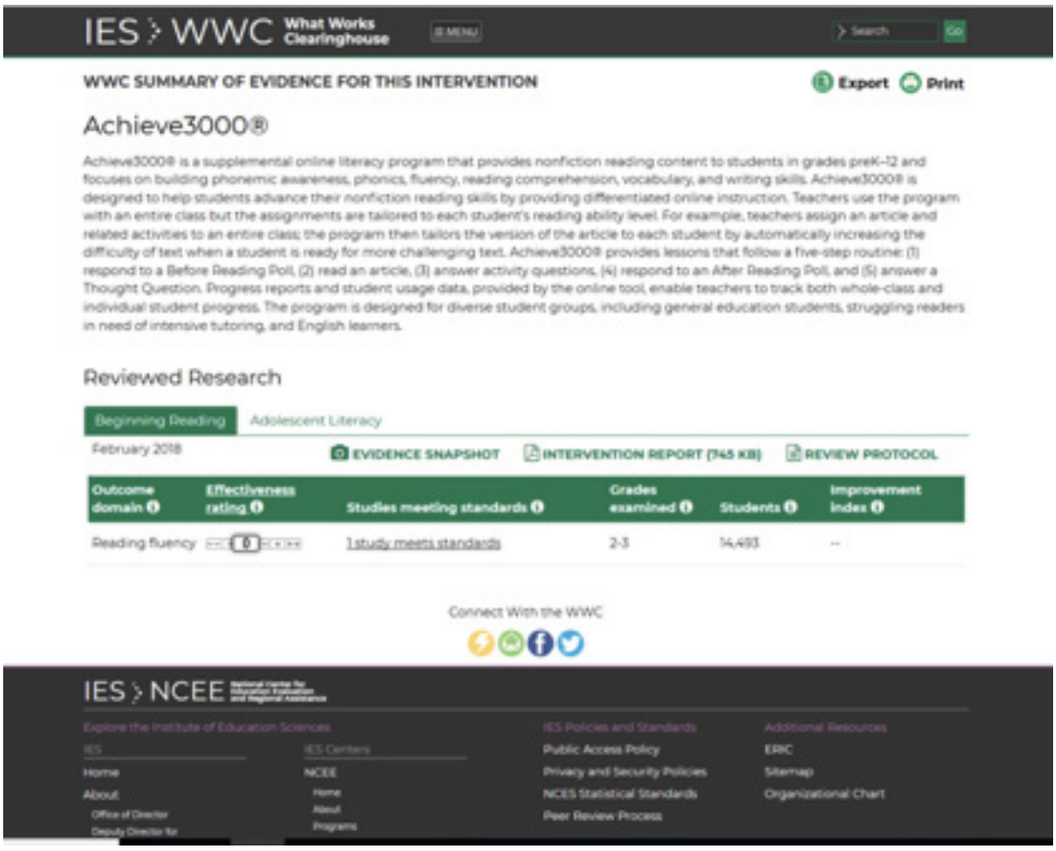
For each intervention there is information on:

- Relevant student population
- Numbers of students studied
- Outcomes measured

- A summary of the effectiveness of an intervention in an outcome domain, based on the quality of research, the statistical significance of findings, the magnitude of findings, and the consistency of findings across studies.
- An indicator of the size of the effect from using the intervention. It is the expected change in percentile rank for an average comparison group student if the student had received the intervention, ranging from -50 to +50.
- The number of studies that met WWC design standards and provide evidence of effectiveness.

Screeenshot

https://ies.ed.gov/ncee/wwc/Intervention/1284



Website pages and searching

Clearinghouse website: https://ies.ed.gov/ncee/wwc/FWW
What we do page lists reports with standards: https://ies.ed.gov/ncee/wwc/WhatWeDo
Procedures handbook v4: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_handbook_v4.pdf
Reviewer Standards handbook v4: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf
Guidance: file:///C:/Users/David/Documents/Study/Homelessness/Data/IES/wwc_reviewer_guidance_103017.pdf

The home page allows searching by topic and student grade level.
The report page provides links to:

- Evidence snapshot
- Full report
- Review protocol
- Relevant Excel files (to export data)

National Institute for Health and Care Excellence (NICE)

Topics

Evidence-based recommendations on a wide range of topics, from preventing and managing specific conditions, improving health and managing medicines in different settings, to providing social care to adults and children, and planning broader services and interventions to improve the health of communities. NICE has produced:

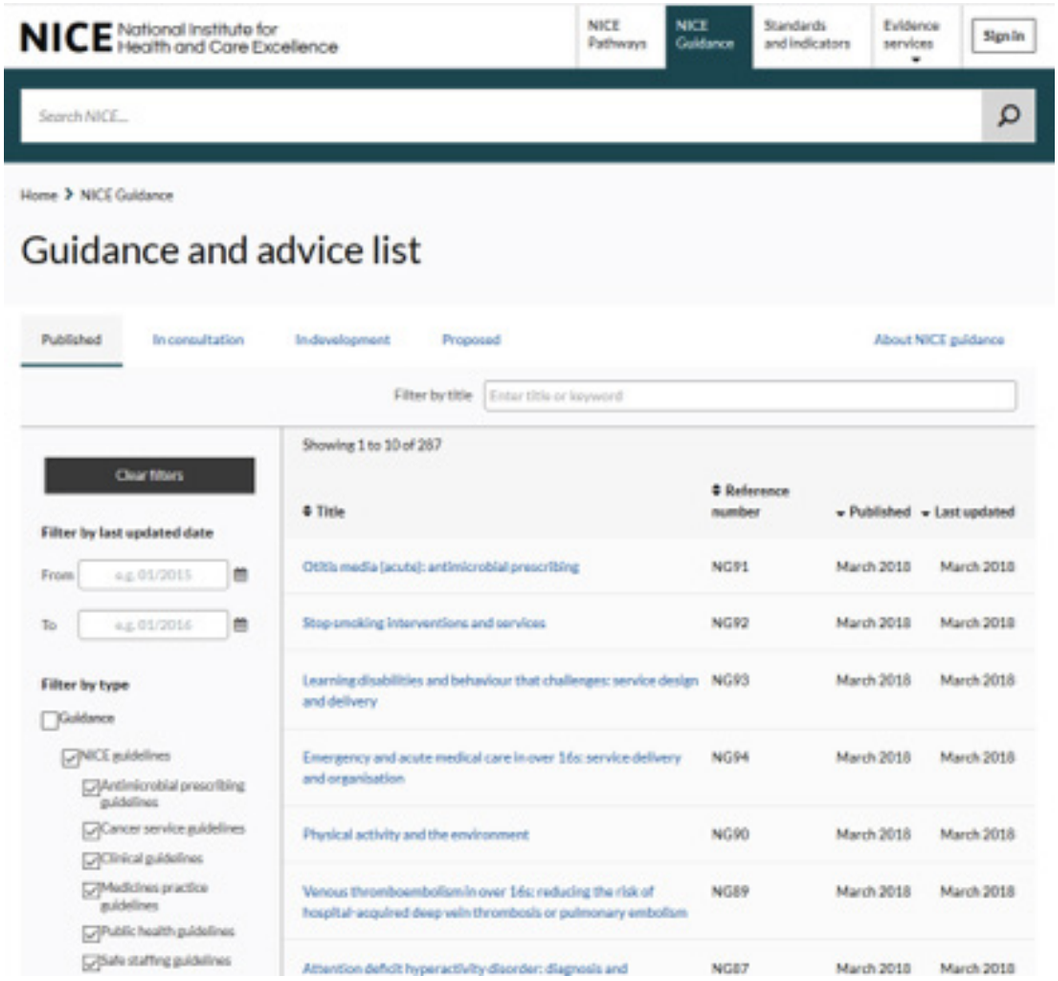
- Clinical guidelines
- Public health guidelines
- Social care guidelines
- Safe staffing guidelines
- Antimicrobial prescribing guidelines
- Medicines practice guidelines
- Cancer service guidelines

Metrics

For each intervention there are Evidence Tables on:

- Critical appraisal, based on appropriate standard checklist, depending on the question
- Research aims: Study aims; Study design; Methodology; Country
- PICO: Population, Intervention, Comparison, Outcomes; Sample
- Findings: Framework; Narrative; Effect sizes

Screeenshot
https://www.nice.org.uk/guidance/published?type=apg,csg,cg,mpg,ph,sg,sc



Website pages and searching

NICE What we do: https://www.nice.org.uk/about/what-we-do
Find Guidance: https://www.nice.org.uk/guidance
Guidelines manual: https://www.nice.org.uk/process/pmg20/chapter/introduction-and-overview#main-stages-of-guideline-development

Example of guidance page (with link to evidence): https://www.nice.org.uk/guidance/ng82/chapter/Recommendations#pharmacological-management-of-amd

Guidance can be searched by type of guidance (as listed under topics above), by date of last update and by category:

- Conditions and diseases
- Health protection
- Lifestyle and wellbeing
- Population groups
- Service delivery, organisation and staffing
- Settings

What Works Centre for Crime Reduction: Crime Reduction Toolkit

Topics

The Crime Reduction Toolkit summarises the best available research evidence on what works to reduce crime. It uses the EMMIE framework (see metrics below) to present evidence from systematic reviews of research on crime reduction interventions in a format that helps users to access and understand it quickly. The toolkits provides evidence on 54 interventions, 52 of which have some evidence of effectiveness.

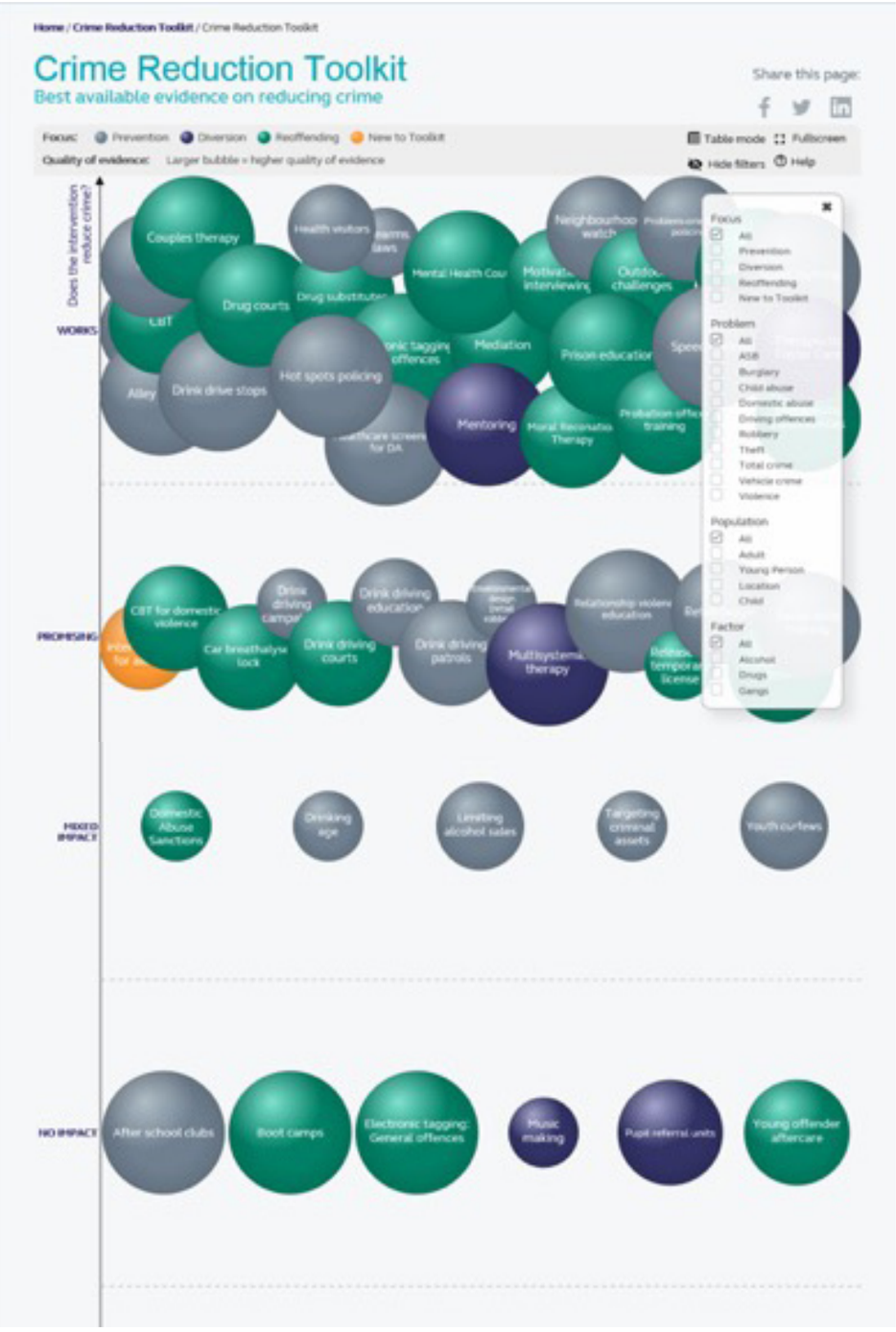
Metrics

For each intervention the toolkit presents the EMMIE metrics of:
Critical appraisal tables

- Impact on crime: Effect
- How it works: Mechanism
- Where it works: Moderator
- How to do it: Implementation
- What it costs: Economic cost

Screeenshot

In 'Bubble mode'. The toolkit is also available in 'Table mode'.
http://whatworks.college.police.uk/toolkit/Pages/BubbleMode.aspx#filter=1



Website pages and searching

Toolkit home page: http://whatworks.college.police.uk/toolkit/Pages/Toolkit.aspx
Reporting guidance: Writing a systematic review for inclusion on the Crime Reduction Toolkit: http://whatworks.college.police.uk/toolkit/About-the-Crime-Reduction-Toolkit/Pages/systematic-reviewers.aspx
Quality scale: http://whatworks.college.police.uk/toolkit/Pages/Quality-Scale.aspx
The home page allows searching by topic and level of evidence of effectiveness. The home page provides links to:

- Effect scale
- Quality Scale
- The 81 interventions within EMMIE

What Works Centre for Local Economic Growth Evidence Reviews and Toolkits

Topics

The What Works Centre for Local Economic Growth analyses which policies are most effective in supporting and increasing local economic growth.

Evidence Reviews consider a specific type of evidence – impact evaluation – that seeks to understand the causal effect of broad policy interventions and to establish their cost-effectiveness. 'Evidence Reviews' are available on:

- Access to Finance
- Apprenticeships
- Area Based Initiatives
- Broadband
- Business Advice

- Employment Training
- Estate Renewal
- Innovation
- Public Realm
- Sport and Culture
- Transport

Toolkits sit beneath the 'Evidence Reviews' and are narrative reviews of types of actions (policy design elements) with a SMS 2 criteria for evidence. Toolkits are available on:

- Mentors
- Public Advisors
- Subsidised Consulting
- Training
- Accelerators
- Incubators
- Investment Promotion Agencies
- Export Promotion Agencies
- Export Credit Agencies

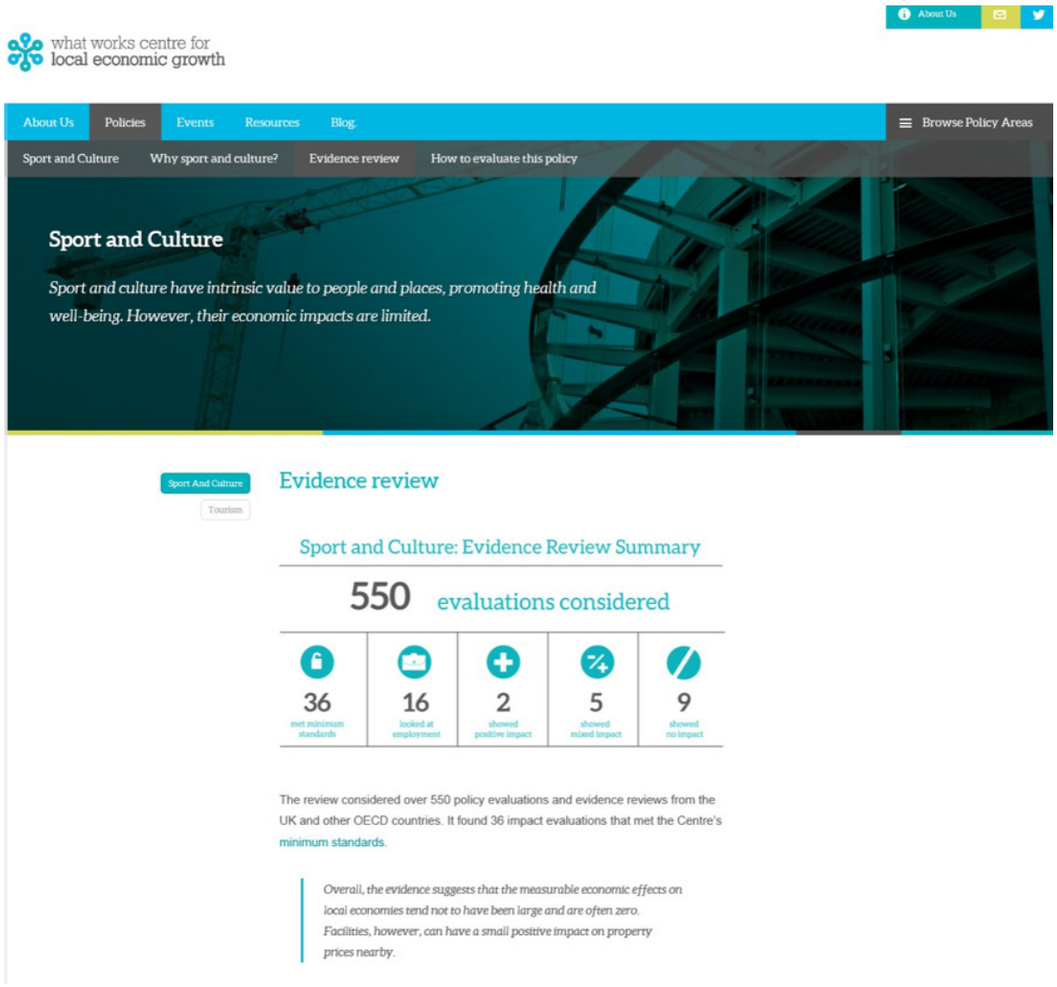
Metrics

For each intervention the toolkit presents the metrics of:

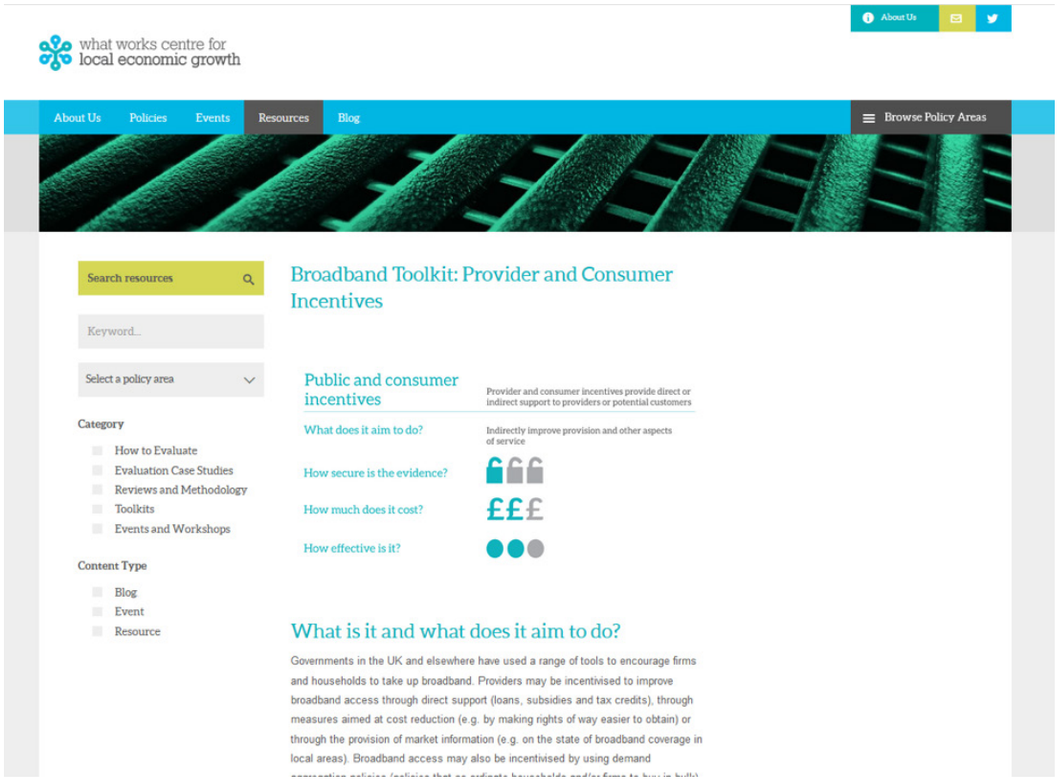
- What do they aim to do?
- How secure is the evidence?
- How much do they cost?
- How effective are they?

Screenshot for What Works Local Economic Growth Evidence Review

http://www.whatworksgrowth.org/policy-reviews/sports-and-culture/evidence-review/



Screenshot for What Works Local Economic Growth Toolkit
http://www.whatworksgrowth.org/resources/broadband-toolkit-provider-and-consumer-incentives-1



Website pages and searching

Resources home page: http://www.whatworksgrowth.org/resources/
Guidance to scoring evidence: http://www.whatworksgrowth.org/resources/scoring-guide/
EG Neigbourhood Saturation Toolkit: http://www.whatworksgrowth.org/resources/toolkit-neighbourhood-saturation-programmes/

The resources page allows searching for Evidence Reviews and Toolkits with filters for policy areas and a key word search.

What Works in Reentry Clearing House

Please note that WWRC pages are being transferred to the National Institute of Justice as part of its CrimeSolutions.gov database. NIJ took the studies in the WWRC and reevaluated them using their rating system, and the ones that met the CS standards can now be found in the Corrections & Reentry section of CS at: https://crimesolutions.gov/TopicDetails.aspx?ID=2.

Topics

The What Works in Reentry Clearinghouse is a "one-stop shop" for research on the effectiveness of a wide variety of reentry programs and practices. To be included in What Works, a study must evaluate how a particular program, practice, or policy affects at least one of a number of relevant outcomes (e.g., recidivism, substance use, housing, employment, or mental health) for people returning to the community from incarceration. Main topic and focus areas are:

- Brand Name Programs
- Case Management and Comprehensive Programs
- Cognitive Behavioural Treatment
- Education
- Employment
- Family-Based Programs
- Housing
- Mental and Physical Health
- Sex Offender Treatment
- Substance Abuse
- Supervision and Sanctions
- Youth Reentry and Aftercare Programs

Metrics

For each intervention there is information on:

Evaluation Rigor:

- High
- Basic

Outcomes:

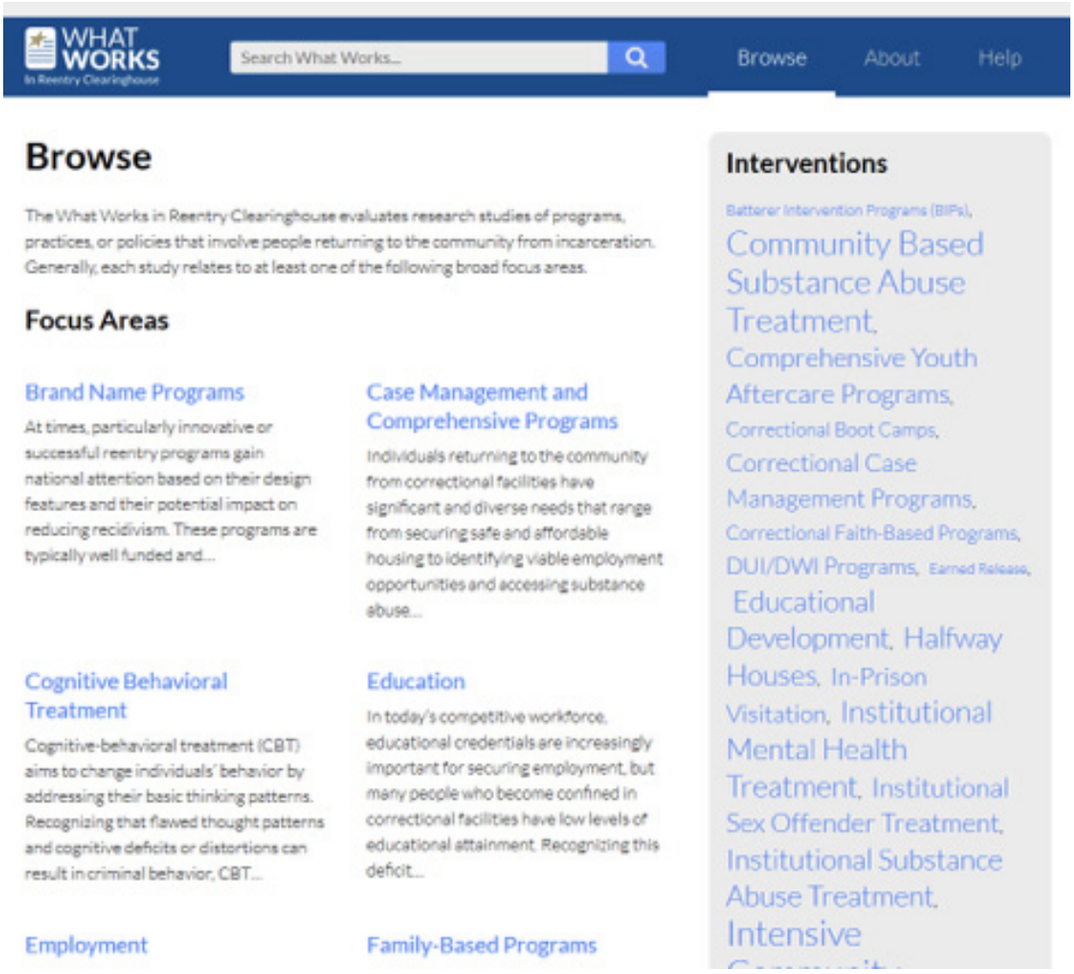
- Strong Beneficial Evidence
- Modest Beneficial Evidence
- No Evidence of Effect
- Modest Harmful Evidence
- Strong Harmful Evidence

For outcome measures of:

- Recidivism
- Employment
- Substance abuse

For each set of studies (with the same Rigor and effectiveness outcomes on the same measures): Findings, Methodology, Methodology limitations, Study population, Quality of implementation.

Screeenshot
https://whatworks.csgjusticecenter.org/browse



Website pages and searching

Clearinghouse website: https://whatworks.csgjusticecenter.org/
About: https://whatworks.csgjusticecenter.org/about
Browse: Focus areas: https://whatworks.csgjusticecenter.org/browse
Example of focus area: https://whatworks.csgjusticecenter.org/focus-area/cognitive-behavioral-treatment

The home page has a generic search function. The home page also has a link to Browse focus areas. This lists the topic and focus areas. Each of these areas provides a summary of Evaluations and Outcomes (Rigor and effectiveness on different outcomes with links to the groups of studies) and links to the different programme types with the

Evaluations and Outcomes (specific studies), Program Description, Recommendations for Practice, and Suggestions for Future Research.

What Works Wellbeing Evidence Comparison Tool

Topics

The What Works Centre provides evidence, guidance and discussion papers on a range of topics relevant to wellbeing. To date the Evidence Comparison Tool has covered a sub-set of this work on wellbeing at work including:

- Wellbeing training
- Organisation-wide approaches
- Changes to ways of working alongside training
- Leadership training
- Training to improve job quality
- Professional training

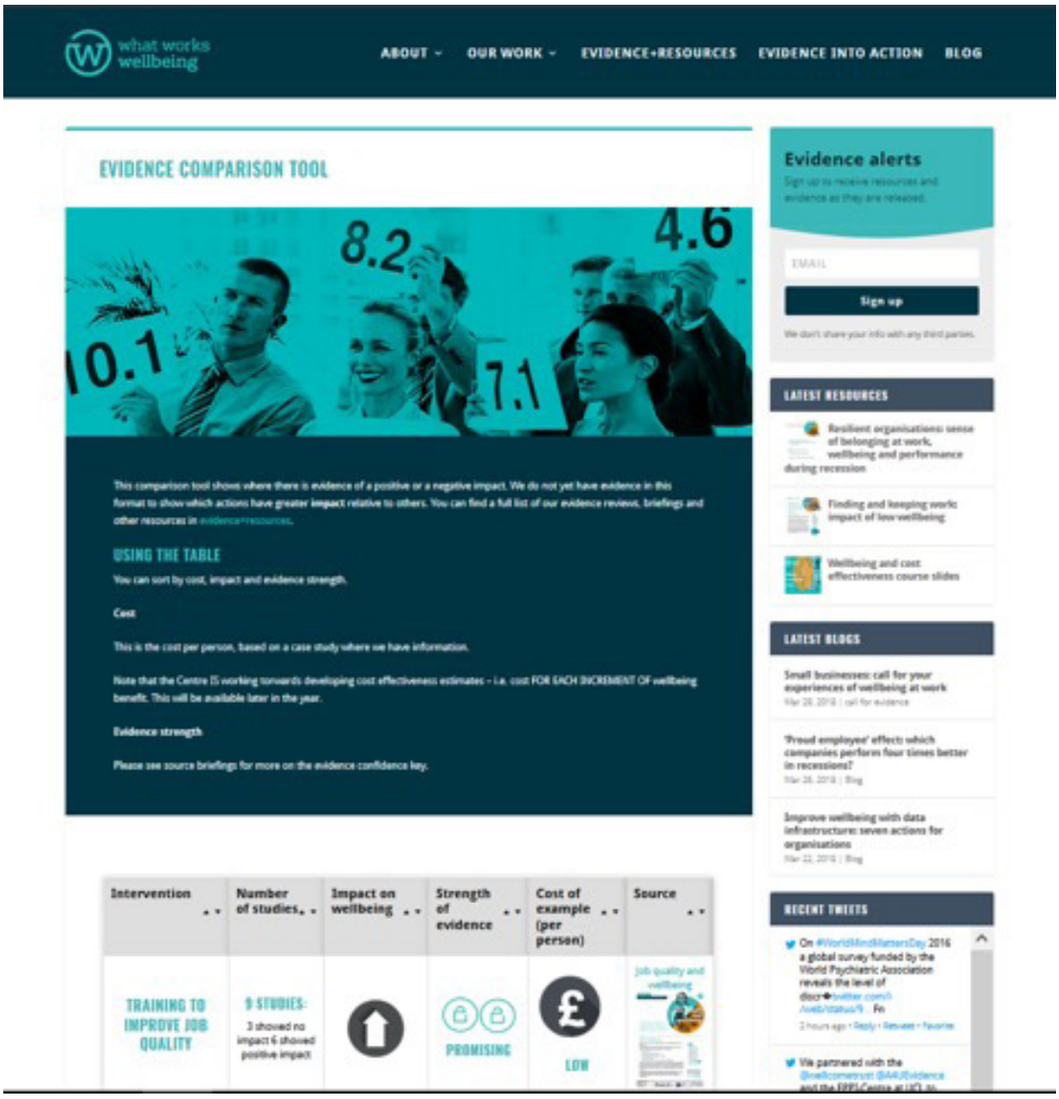
The comparisons are from evidence provided by two evidence briefings on: Learning at work and wellbeing; and Job quality and wellbeing.

Metrics

For each intervention the toolkit presents the metrics of:

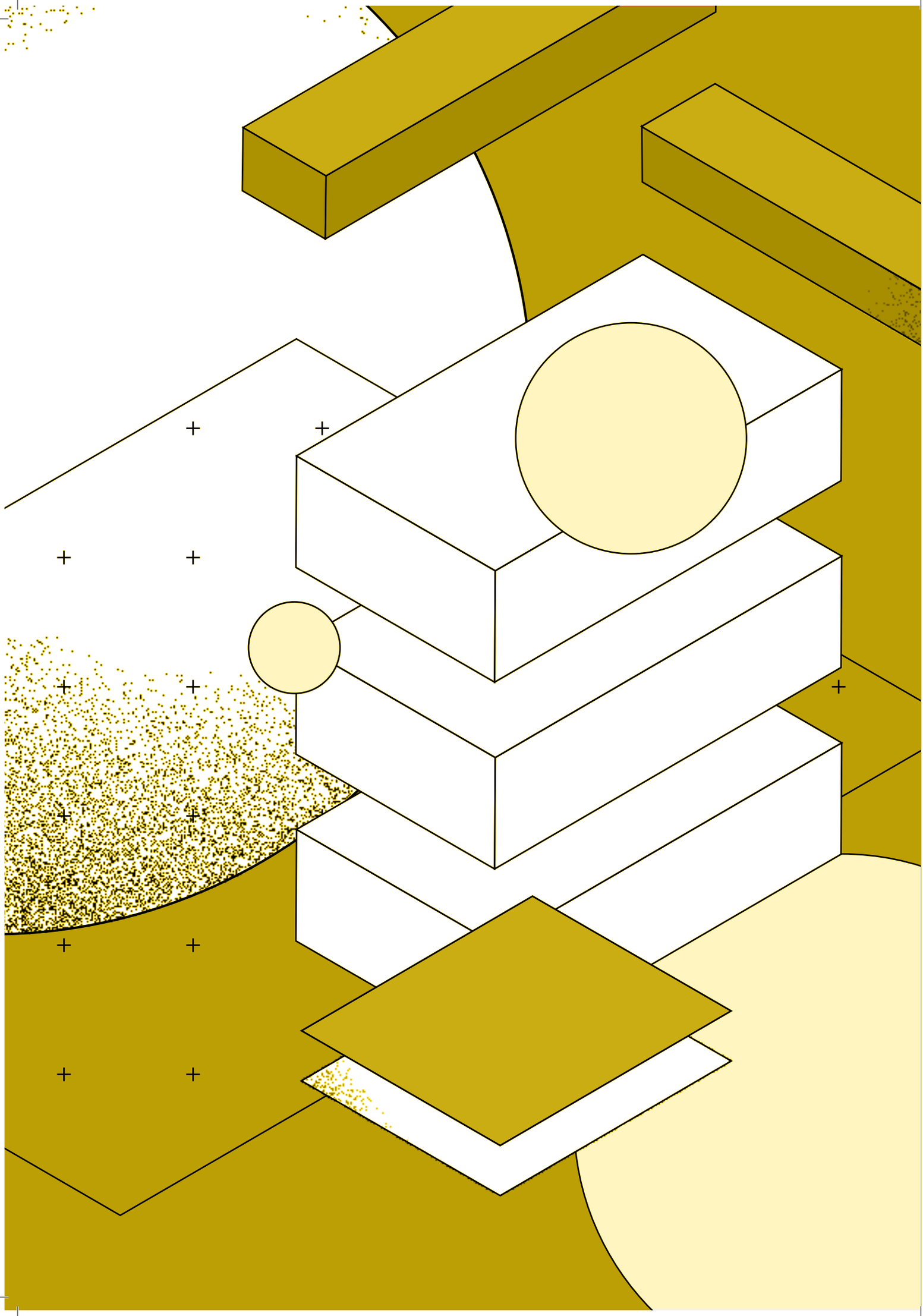
- Intervention
- Number of studies (and whether they reported a positive, negative or no impact)
- Impact on wellbeing (direction of)
- Strength of evidence
- Cost of example (per person)
- Source (of evidence)

Screenshot for What Works Wellbeing
https://www.whatworkswellbeing.org/evidence-comparison-tool/



Website pages and searching
Evidence review methods: https://www.whatworkswellbeing.org/product/a-guide-to-our-evidence-review-methods/

Currently all of the evidence comparisons can be seen on the main Evidence Comparison Tool webpage.



Appendix 2

Main components of statements about the impact evidence base

California Evidence-Based Clearinghouse for Child Welfare (CEBC): EFFECTIVE PROGRAMMES

Evidence base review method

(Technical method of review + Relevance of method + Relevance of focus)

REVIEW APPROACH / EVIDENCE STANDARDS

(Technical method + Relevance of method + Relevance of focus)

Examination of available research evidence. Seeking multiple rigorous primary studies showing positive impact and no harm.

Typically, the raters include a topic expert and two CEBC staff. If there is a discrepancy, the CEBC Scientific Director makes the final decision.

Main focus of the portal's research questions:
Effectiveness and implementation of programmes.

REVIEW INCLUSION CRITERIA

Study method: 1. Primary evaluation studies: randomized controlled trial, or utilizing some form of control (e.g., untreated group, placebo group, matched wait list study). The study has been reported in published, peer-reviewed literature. Outcome measures must be reliable and valid, and administered consistently and accurately across all subjects.

Evidence claim evidence standards

(Method of review + Included Studies + Nature and extent of totality of evidence)

STRENGTH OF EVIDENCE / BENEFIT

Effectiveness rating: 5 point CEBC Scientific Rating Scale based on the criteria of a certain number of studies both [(i)] showing an effect; and [(ii)] meeting evidence standards:

1. Well-Supported by Research Evidence
2. Supported by Research Evidence
3. Promising Research Evidence
4. Evidence Fails to Demonstrate Effect
5. Concerning Practice
NR = Not able to be Rated on the CEBC Scale

Highest rating 1 = Multiple Site Replication and Follow-up:

At least two rigorous randomized controlled trials (RCTs) in different usual care or practice settings have found the practice to be superior to an appropriate comparison practice. In at least one of these RCTs, the practice has shown to have a sustained effect at least one year beyond the end of treatment, when compared to a control group. The RCTs have been reported in published, peer-reviewed literature. Outcome measures must be reliable and valid, and

Guidance

Details on nature of programmes provided to inform decisions on choice of programme and implementation (see Platform Descriptions).

Not specific guidance on action but guidance on how to identify needs, select programmes, and implement them.

EPIS (Exploration, Preparation, Implementation, Sustainment) framework on stages of implementation: http://www.cebc4cw.org/implementing-programs/tools/epis/
And an implementation guide: http://www.cebc4cw.org/implementing-programs/guide/

Other criteria: On topic. The focus is on the efficacy of programmes.

STUDY IDENTIFICATION

In addition to the information provided by the program, the CEBC staff conduct a literature search on each program to obtain any published, peer reviewed research. See Review and rating process: http://www.cebc4cw.org/home/how-are-programs-on-the-cebc-reviewed/

INCLUDED STUDY EVIDENCE STANDARDS

(Technical method + Relevance of method + Relevance of focus)

Same criteria about primary studies as specified in rating the effectiveness of intervention programmes (Review evidence standards).

SPECIFICATION OF PORTAL'S METHODS

Report specific on portal's studies: -

Manuals on portals evidence base methods: -

Manuals on study appraisal: Scientific Rating Scale.

Explanation of criteria for appraisal specifically: -

Internal quality assurance: The CEBC Scientific Director makes the final decision.

administered consistently and accurately across all subjects. If multiple outcome studies have been published, the overall weight of the evidence supports the benefit of the practice.

There is no case data suggesting a risk of harm that: a) was probably caused by the treatment and b) the harm was severe or frequent. There is no legal or empirical basis suggesting that, compared to its likely benefits, the practice constitutes a risk of harm to those receiving it. The practice has a book, manual, and/or other available writings that specify components of the service and describe how to administer it.

Measurement Tools Rating Scale: describes the CEBC ratings for tools used for screening or assessment, is a three-level rating (A, B, or C) scale based on the level of psychometrics (e.g., sensitivity and specificity, reliability and validity) found in published, peer-reviewed journals.

IMPACT (IF SEPARATELY CONSIDERED) -

NATURE AND TOTALITY OF EVIDENCE

Focus is on totality of evidence. Extent of evidence is through numbers of studies.

PROCESS

COSTS

The nature of resources required to run the programmes are listed.

Clearinghouse for Labor Evaluation and Research (CLEAR)

Evidence base review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* CLEAR identifies and reviews relevant research and, for causal studies that estimate impact, provides ratings according to the quality of the evidence they produced. CLEARprovides access to the reviews of the research evidence from each of these studies on its portal. These are expert systematic narrative reviews of studies (using specified standards depending on the type of research design). A number of syntheses of the research are also provided. The process is: 1.Identify research questions of interest to DOL program administrators and policymakers that may be appropriate for systematic evidence reviews in relevant topic areas. 2. 2. Work with content experts to conduct initial feasibility assessments on the size and type of the literature base to answer those questions. If appropriate, continue to work with content experts to develop transparent, publicly available protocol for conducting systematic evidence reviews to answer the research questions within the topic areas. Using the parameters transparently outlined	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT The portal predominantly provides appraisals of individual studies (see included study evidence standards). There are some syntheses for which a view is taken across the breadth of studies meeting the individual study evidence standards. IMPACT (IF SEPARATELY CONSIDERED) – NATURE AND TOTALITY OF EVIDENCE - PROCESS – COSTS-	The website provides information studies that examine the implementation of interventions. Profiles of research describing implementation experiences include some key considerations for interpreting the findings. This can help inform to what extent the findings from a study might apply to particular situations and other expert insights and cautions. Stakeholders are also involved in specifying the research questions that lead to topic areas for study.

in the protocol (including specific keywords, research databases and dates), CLEAR searches the literature for research addressing those questions that meet protocol standards.

3. Review and, if appropriate, rate these studies and create profiles capturing the main features of each report or journal article.

4. Produce more detailed profiles and ratings for causal research estimating impact.

5. Develop and maintain a searchable database of all the relevant research (profile summaries and citations) identified.

6. Synthesize the research across studies within a topic area, highlight gaps in the literature, and suggest areas in which further research is needed.

Main focus of the portal's research questions: Effectiveness and implementation of whatever actions evaluated within portal topic areas (includes programmes and wider approaches).

REVIEW INCLUSION CRITERIA:

Study method: Primary evaluation studies: randomized controlled trial and other causal designs (including nonexperimental designs).

Depending on the topic, CLEAR also reviews descriptive or implementation studies.

Other criteria: On topic.

STUDY IDENTIFICATION

Topic specific search strategies, according to each topic area's protocol.

INCLUDED STUDY EVIDENCE STANDARDS

(Technical method + Relevance of method + Relevance of focus)

This is only about the methodological standards and the causal impacts estimated (the study's ability to estimate causal impacts), whether positive, negative, or null. It is not a scale of effectiveness. The results of the studies need to be considered alongside the rating of their evidence standards.

High causal evidence: There is strong evidence that the estimated effects are solely attributable to the program or policy being examined. This rating can apply only to RCTs and ITS designs.

Moderate causal evidence: There is moderate evidence that the estimated effects are attributable at least in part to the program or policy being examined. However, there might be other factors that were not accounted for and that might also have contributed to the estimated effects. This rating can apply to nonexperimental designs. It can also apply to RCTs and ITS designs that do not meet the criteria for a high causal evidence rating.

Low causal evidence: There is little evidence that the estimated effects are attributable solely to the intervention; other factors are likely to have contributed. Applies to all causal designs that do not meet the criteria for high or moderate causal evidence

ratings.

For more details see Clear Causal Evidence Guidelines, Version 2.1. For implementation studies, see Guidelines for Reviewing Implementation Studies

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: Topic-specific protocols, policies and procedures.

Manuals on study appraisal: CLEAR causal evidence guidelines, study review guides (SRGs) Causal Evidence Guidelines, Version 2.1 (on primary studies).

Explanation of criteria for appraisal specifically:

CONSERVATION EVIDENCE

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* Evidence is then collated and then panels of experts assess the collated evidence for each intervention to determine its effectiveness, the certainty of the evidence and, in most cases, whether there are negative side-effects on the group of species or habitat of concern (harms). Report called a synopsis. **Main focus of the portal's research questions:** the effectiveness of specific actions (one synopsis considers the evidence for all/many actions within a topic area, usually a species group or habitat). REVIEW INCLUSION CRITERIA: **Study method:** Only studies that have quantitatively monitored the effect of an action are included in a synopsis. Predictive modelling studies and studies looking at species distributions in areas with long-standing management histories (correlative studies) are excluded. **Other criteria:** On topic. STUDY IDENTIFICATION: General and specific journals are systematically searched for evidence on all actions to create a subject-wide literature	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT Using these assessments, interventions are categorised, based on a combination of **effectiveness** (the size of benefit or harm) and **certainty** (the strength of the evidence). **Effectiveness:** 0 = no effect, 100% = always effective. **Certainty of the evidence:** 0 = no evidence, 100% = high quality evidence; complete certainty. This is certainty of effectiveness of intervention, not of harms. **Harms:** 0 = none, 100% = major negative side-effects to the group of species/habitat of concern. The median score from all the experts' assessments is calculated for the effectiveness, certainty and harms for each intervention. Effectiveness categorization is based on these median values, i.e. on a combination of the size of the benefit and harm and the strength of the evidence. The following categories are used: • Beneficial • Likely to be beneficial • Trade-off between benefit and harms • Unknown effectiveness • Unlikely to be beneficial	Video on combining experience and local knowledge with social values and evidence.

database. The database and searches of additional /updates of specific journals provides studies used for each synopsis.

INCLUDED STUDY EVIDENCE STANDARDS

(Technical method + Relevance of method + Relevance of focus)

Assessed by a panel of experts.

SPECIFICATION OF METHOD

Report specific on portal's studies: Manuals on portal's evidence base methods: video providing a basic overview of the process for producing a synopsis of evidence in the style of Conservation Evidence.

Manuals on study appraisal: Explanation of criteria for appraisal specifically: Explained on web page at: https://www.conservationevidence.com/content/page/79

• Likely to be ineffective or harmful.

IMPACT (IF SEPARATELY CONSIDERED)

The size of the benefit and harm used to assess likely benefit.

NATURE AND TOTALITY OF EVIDENCE

The studies included and their findings for each action is provided.

PROCESS

Some process information may be provided in background information in the synopses.

COSTS

-

EARLY INTERVENTION FOUNDATION: EIF GUIDEBOOK

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* Two main methods of identifying research evidence: 1 –Large-scale reviews of programmes within a given thematic group (e.g. parenting programmes): commissioned systematic reviews using main bibliographic databases, grey literature, other databases. 2 – For smaller-scale reviews on individual programmes: Provider submitted impact studies plus additional web-based search conducted by EIF researchers to identify other potentially relevant evaluations and studies. Examination of available research evidence using an interpretative approach informed by rapid realist reviews and qualitative synthesis. Data is extracted using a three-stage process of: • Reading all included papers to identify key features. • Producing narrative reviews of the findings/issues from groups of papers addressing specific questions. • Re-reading key papers to ensure review findings and arguments are supported by research.	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT **Evidence effectiveness** ('Evidence Standard') 4+ point scale based on the criteria of a certain number of studies both showing (i) an effect; and (ii) meeting evidence standards: **Level 4+:** The criteria for level 4 plus: At least one of the effectiveness evaluations will have been conducted independently of the programme developer. The intervention has evidence of improving EIF child outcomes from three or more rigorously conducted evaluations (RCT/ QED) conducted within real world settings. **Level 4:** The programme has evidence from at least two rigorously conducted evaluations (RCT/QED) meeting Level 3 criteria and: At least one evaluation uses a form of measurement that is independent of the study participants (and also independent of those who deliver the programme). At least one evaluation has evidence of a long-term outcome of 12 months. **Level 3+:** Meets Level 3 criteria and: Additional consistent positive evidence from other	The evidence ratings of programmes are not a 'kite mark' or guarantee of effectiveness or relevance to in local contexts. They are not recommendations for programmes to be selected 'off the shelf'. They are instead a starting point for finding out more about effective early interventions that might be relevant in specific contexts.

Main focus of the portal's research questions: the appraisal of the effectiveness of particular programmes.

REVIEW INCLUSION CRITERIA:

Study method: Primary evaluation studies: Randomized controlled trial, quasi-experimental design, pre-post designs.

Other criteria: On topic - a piece of research investigating the effect of an identified early intervention programme.

STUDY IDENTIFICATION:
-
INCLUDED STUDY EVIDENCE STANDARDS

(Technical method + Relevance of method + Relevance of focus)

Same criteria about primary studies as specified in rating the effectiveness of intervention programmes (Review evidence standards).

There is a formal process within which evidence standards judgements are made:

• Initial assessment: Using this material, EIF conducts an initial assessment against 33 separate criteria relating to the quality and rigour of the evaluations that provide a programme's best evidence. For more detail on the standards of evidence, see: EIF evidence standards.

• Expert review: EIF's initial assessment is reviewed by a panel of academics and experts with knowledge of the specific subject area and of evaluation and statistical analysis.

• Sub-panel reviews: Small groups of subject-matter experts and EIF evaluations (occurring under ideal circumstances or real world settings) that do not meet this criteria, thus keeping it from receiving an assessment of 4 or higher.

Level 3: Evidence from at least one evidence or higher is the point at which there is sufficient confidence that a causal relationship can be assumed and:

Participants are randomly assigned to the treatment and control groups through the use of methods appropriate for the circumstances and target population, OR sufficiently rigorous quasi-experimental methods (e.g. regression discontinuity, propensity score matching) are used to generate an appropriately comparable sample through non-random methods.

IMPACT (IF SEPARATELY CONSIDERED)

NATURE AND TOTALITY OF EVIDENCE

Number of primary studies meeting criteria (rather than the extent of evidence within each study)

PROCESS

Some information is provided on the theory of change of the intervention (by what mechanisms it is hypothesized to have an effect).

COSTS

An assessment of the relative input costs of early intervention programmes, such as practitioners' and supervisors' time, qualifications or training requirements. Rated on a 5 point

staff meet to review the detail of each assessment in fine detail and agree a provisional evidence rating. The makeup of the panel changes with each panel, to ensure the right mix of expertise, are included.

• Confirming the rating: Provisional evidence and cost ratings are shared with providers of the interventions, who may request a reassessment if they consider that the criteria have not been properly applied. Following this stage, a final moderation meeting is held with all members of the sub-panel meetings to ensure consistency of rating and ratify provisional ratings as final.

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: -

Manuals on study appraisal:

Explanation of criteria for appraisal specifically: EIF Evidence Standards.

Internal quality assurance: Stages for different staff, experts, and programme developers to consider the appraisal of primary studies.

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: -

Manuals on study appraisal: -

Explanation of criteria for appraisal specifically: EIF Evidence Standards.

scale.
Highest rating of 5 High Cost: more than £2,000 per recipient cost based on the set up and inputs required (time, no. of families to be served, staff qualifications, fees).

Internal quality assurance:
Stages for different staff, experts, and programme developers to consider the appraisal of primary studies.

Evidence Strength: EEF 'padlocks' evidence ratings

EDUCATION ENDOWMENT FOUNDATION (EEF)

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* Review of reviews of existing systematic statistical meta analysis reviews plus primary impact studies where relevant. **Main focus of the portal's research questions:** Types of approach and programmes. REVIEW INCLUSION CRITERIA **Study method:** In order of preference: 1. Existing meta analyses of experimental evaluations; 2.exisitng syntheses of less controlled evaluations; 3. Any individual studies experimental evaluations plus single studies (If not available, then observational data, then individual studies). **Other criteria:** On topic. STUDY IDENTIFICATION - INCLUDED STUDY EVIDENCE STANDARDS *(Technical method + Relevance of method + Relevance of focus)* SPECIFICATION OF METHOD **Report specific on portal's studies:** -	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT Evidence Strength: EEF 'padlocks' evidence rating The Toolkits presents a 5 point rating of the security of the evidence underpinning each topic based on: (i) the quantity of evidence available (i.e. the number of systematic reviews and the number of primary studies which they synthesise); (ii) the methodological quality of the available evidence; and (iii) the consistency of estimated impact across the systematic reviews and meta-analyses that have been synthesised. **= Very limited:** Quantitative evidence of impact from single studies, but with effect size data reported or calculable. No systematic reviews with quantitative data or meta-analyses located. **= Limited:** At least one meta-analysis or systematic review with quantitative evidence of impact on attainment or cognitive or curriculum outcome measures. **= Moderate:** Two or more rigorous meta-analyses of experimental studies of school age students with cognitive or curriculum outcome measures.	Guidance reports for practitioners using toolkit evidence of efficacy plus a wide range of robust evidence from other studies and reviews. A scoping document setting out headline recommendations and supporting evidence is revised with support and feedback from an advisory panel of teachers and researchers.

Manuals on portals evidence base methods: Technical appendix and process manual (Working document v.01) July 2018.

Manuals on study appraisal: -

Explanation of criteria for appraisal specifically: -

= Extensive: Three or more meta-analyses from well-controlled experiments mainly undertaken in schools using pupil attainment data with some exploration of causes of any identified heterogeneity.

= Very Extensive: Consistent high quality evidence from at least five robust and recent meta-analyses where the majority of the included studies have good ecological validity and where the outcome measures include curriculum measures or standardised tests in school subject areas.

IMPACT (IF SEPARATELY CONSIDERED)

Effect sizes scaled as months of additional progress increased (or decreased) taking average pupil progress over a year as a benchmark. For the Early Childhood Checklist the outcome is cognitive development.

1 month progress = up to 0.02 ES;
12 months = up to 0.96 ES

NATURE AND TOTALITY OF EVIDENCE

Built into effectiveness rating scale.

PROCESS

-

COSTS

Costs of adoption including training of staff. 5 point scale:

£: Very low: up to £2,000 per year per class of 25 pupils, or less than £80 per pupil per year.

££: Low: £2,001 to £5,000 per

year per class of 25 pupils, or up to £200 per pupil per year.

£££: Moderate: £5,001 to £18,000 per year per class of 25 pupils, or up to £720 per pupil per year.

££££: High: £18,001 to £30,000 per year per class of 25 pupils, or up to £1,200 per pupil.

£££££: Very high: over £30,000 per year per class of 25 pupils, or over £1,200 per pupil.

EUROPEAN MONITORING CENTRE FOR DRUGS AND DRUG ADDICTION (EMCDDA): BEST PRACTICE PORTAL EVIDENCE DATABASE

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* Examination of available existing research reviews of the literature. **Main focus of the portal's research questions:** Effectiveness of different approaches. REVIEW INCLUSION CRITERIA: **Study method:** Existing research reviews of the literature. **Other criteria:** - STUDY IDENTIFICATION: Listing of search sources and common search terms. INCLUDED STUDY EVIDENCE STANDARDS *(Technical method + Relevance of method + Relevance of focus)* SPECIFICATION OF METHOD **Report specific on portal's studies:** - **Manuals on portals evidence base methods:** Evidence ratings and search approach at: http://www.emcdda.europa.eu/best-practice/evidence/about	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT **Evidence ratings** 5 point scale based on BMJ Evidence and previous work by the Cochrane Collaboration. Scores made on basis of existing narrative and systematic reviews of research. The available information on the effects of specific interventions are examined and then ranked as: **Beneficial:** Interventions for which precise measures of the effects in favour of the intervention were found in the systematic reviews of randomised controlled trials (RCTs), and that were recommended in guidelines with reliable methods for assessing evidence (such as GRADE*). An intervention ranked as 'beneficial' is suitable for most contexts. **Likely to be beneficial:** Interventions that were shown to have limited measures of effect, that are likely to be effective but for which evidence is limited, and/or those that are recommended with some caution in guidelines with reliable methods for assessing evidence (such as GRADE). An intervention ranked as 'likely to be beneficial'	Policy and practice briefings provide: an overview of an issues including potential responses: a summary of the available evidence, and the implications for policy and practice.

Manuals on study appraisal: -

Explanation of criteria for appraisal specifically: Best practice portal evidence ratings.

is suitable for most contexts, with some discretion.

Trade-off between benefits and harms: Interventions that obtained measures of effects in favour of harm reduction and/or are recommended in guidelines with reliable methods for assessing evidence (such as GRADE), but that showed some limitations or unintended effects that need to be assessed before providing them.

Unknown effectiveness: Interventions for which there are not enough studies or where available studies are of low quality (with few patients or with uncertain methodological rigour), making it difficult to assess if they are effective or not. Interventions for which more research should be undertaken are also grouped in this category. Evidence of ineffectiveness: Interventions that gave negative results if compared with a standard intervention, for example.

IMPACT (IF SEPARATELY CONSIDERED)

-

NATURE AND TOTALITY OF EVIDENCE

-

PROCESS

The wider website includes information on guidelines and best practice which may include process data.

COSTS

-

EUROPEAN PLATFORM FOR INVESTING IN CHILDREN (EPIC): EVIDENCE BASED PRACTICES

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* **Systematic narrative reviews** **Main focus of the portal's research questions:** Effectiveness of programmes REVIEW INCLUSION CRITERIA: **Study method:** Evaluation with at least a comparison group (though +rating for evidence effectiveness with pre/post test can lead to a score of Best Practice) **Other criteria:** The practice was implemented and evaluated in one of the 28 EU Member States. STUDY IDENTIFICATION: Comprehensive literature search INCLUDED STUDY EVIDENCE STANDARDS *(Technical method + Relevance of method + Relevance of focus)* Same criteria about primary studies as specified in rating the effectiveness of intervention programmes (evidence claim evidence standards).	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT Criteria to determine the evidence level are organised according to three categories: **Evidence of effectiveness:** + (pre/post test with sample size of at least20) or ++ (experimental design). **Transferability:** + (has been evaluated in at least one additional population beyond the original study population); ++ (and as been found to be cost-effective/cost-beneficial). **Enduring impact:** Continues to find positive impact at 2 years follow up. These are used for a scale of three evidence levels: • Emergent Practice: An "emergent practice" has achieved at least a + in "evidence of effectiveness." • Promising Practice: A "promising practice" has achieved at least a + in "evidence of effectiveness" and a + in at least one of the other two categories, "transferability" and "enduring impact." • Best Practice: A "best practice" has achieved at least a + in	

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: -

Manuals on study appraisal: - Explanation of criteria for appraisal specifically: Web pages on process ,methods and criteria.

each of the three evidence categories, including "evidence of effectiveness", "transferability" and "enduring impact."

IMPACT

-

NATURE AND TOTALITY OF EVIDENCE

The studies included in assessment of effectiveness are listed.

PROCESS

Information on moderators and mediators may be included under Issues to Consider on evidence page. Measure of transferability provides some indication of extent of evidence.

COSTS

Cost data may be included under Resource Allocation on evidence page. Costs part of transferability measure.

Evidence Based Teen Pregnancy Programs

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* Appraial of causal inference in primary evaluation studies and making overall narrative assessment of findings. **Main focus of the portal's research questions:** Programmes. REVIEW INCLUSION CRITERIA **Study method:** Primary evaluation studies: randomized controlled trial, quasi-experimental design. **Other criteria:** On topic. At least one study that was conducted within the last 20 years. STUDY IDENTIFICATION A broad literature search that includes both published and unpublished work. INCLUDED STUDY EVIDENCE STANDARDS *(Technical method + Relevance of method + Relevance of focus)* A quality rating of high, moderate, or low according to the risk of bias in the study's impact findings: High quality randomized trial; Moderate quality	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT At least one impact study showing evidence of a favorable, statistically significant impact on at least one outcome measure within one of the eligible outcome domains, for either the full analytic sample or a subgroup defined by (1) gender or (2) sexual experience at baseline. The eligible outcome domains are (1) sexual activity; (2) number of sexual partners; (3) contraceptive use; (4) STIs or HIV; and (5) pregnancies. In addition, the study cannot show evidence of any adverse, statistically significant impacts on any outcomes in these domains. For each domain in each program, the evidence of effectiveness is classified as falling into one of the following four categories: **1. Positive impacts:** Evidence of uniformly favorable impacts across one or more outcome measures, analytic samples (full sample or subgroups), and/or studies. **2. Mixed impacts:** Evidence of a mix of favorable, null, and/ or adverse impacts across one or more outcome measures,	

randomized trial.; Moderate quality quasi experimental trial; Low study rating; Did not meet eligibility criteria.

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: Review Protocol Version 5.0.

Manuals on study appraisal: -
Explanation of criteria for appraisal specifically: -

analytic samples (full sample or subgroups), and/or studies.

3. Indeterminate impacts: Evidence of uniformly null impacts across one or more outcome measures, analytic samples (full sample or subgroups), and/or studies.

4. Negative impacts: Evidence of uniformly adverse impacts across one or more outcome measures, analytic samples (full sample or subgroups), and/or studies.

The review team makes these assessments separately for each of the five outcome domains. As a result, a program may be classified as having "positive impacts" in one domain but "mixed impacts" in another domain. In addition, programs are classified in these categories only for the domains on which they have been evaluated. For example, if a program has been evaluated for impacts on sexual activity but not pregnancy, the review team classifies the program's evidence of effectiveness only for the domain of sexual activity. When comparing findings across multiple studies of the same program, the review team bases this comparison whenever possible on the estimated effect sizes and confidence intervals reported in the individual studies.

IMPACT (IF SEPARATELY CONSIDERED)

-

NATURE AND TOTALITY OF EVIDENCE

List of studies meeting evidence standards.

PROCESS

For programs meeting the review criteria for evidence of effectiveness, the review team calculates an implementation readiness score based on: (1) curriculum and materials (2) training and staff support and (3) fidelity monitoring tools and resources.

Information provided on implementation and advice from the field.

A defined logic model is part of the 3rd component of an implementation readiness score.

COSTS

The website provides general tips on Implementation costs to keep in mind.

IES: WHAT WORKS CLEARINGHOUSE

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* Evaluating causal inference in individual studies. **Main focus of the portal's research questions:** Effectiveness of programmes. REVIEW INCLUSION CRITERIA: **Study method:** Primary evaluation studies: randomized controlled trial, quasi-experimental design, regression discontinuity design, and single-case design. **Other criteria:** - STUDY IDENTIFICATION Systematic search INCLUDED STUDY EVIDENCE STANDARDS *(Technical method + Relevance of method + Relevance of focus)* Meets Standards (see Standards Handbook): 1. Meets standards without reservations (only RCTs with low attrition). 2. Meets standards with reservations. 3. Does not meet standards.	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT Effectiveness rating: 6 point scale (Table IV.3 in Procedures Handbook) based on the criteria of a certain number of studies both [i] showing an effect; and [ii] meeting evidence standards for each outcome domain. **Positive effects:** Strong evidence of a positive effect with no overriding contrary evidence with no overriding contrary evidence: Two or more studies show statistically significant positive effects, at least one of which meets design standards without reservations, AND No studies show statistically significant or substantively important negative effects. **Potentially positive effects:** Evidence of a positive effect with no overriding contrary evidence. At least one study shows statistically significant or substantively important positive effects, AND Fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects, AND No studies show statistically significant or substantively important negative effects.	Practice guides based on Strong Evidence, Moderate Evidence or Minimal Evidence Base. Criteria for Strong Evidence: **Validity:** The research has high internal validity and high external validity based on studies that meet standards **Effects on relevant outcomes:** The research shows consistent positive effects without contradictory evidence in studies with high internal validity. Relevance to scope: The research has direct relevance to scope— relevant context, sample, comparison, and outcomes evaluated. **Relationship between research and recommendations:** Direct test of the recommendation in the studies, or the recommendation is a major component of the intervention tested in the studies. **Panel confidence:** Panel has a high degree of confidence that this practice is effective. **Role of expert opinion:** Not applicable **When assessment is the focus of the recommendation:** Assessments meet the standards of The Standards for Educational and Psychological Testing

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: Procedures Handbook V4 + Standards Handbook V4 for reviewing included primary studies + Reviewers guidance handbook.

Manuals on study appraisal: -

Explanation of criteria for appraisal specifically: -

No discernible effects: No affirmative evidence of effects. None of the studies show statistically significant or substantively important effects, either positive or negative.

Mixed effects: Evidence of inconsistent effects. EITHER both of the following: At least one study shows statistically significant or substantively important positive effects, AND At least one study shows statistically significant or substantively important negative effects, BUT no more such studies than the number showing statistically significant or substantively important positive effects. OR both of the following: At least one study shows statistically significant or substantively important effects, AND More studies show an indeterminate effect than show statistically significant or substantively important effects.

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence. EITHER both of the following: One study shows statistically significant or substantively important negative effects, AND No studies show statistically significant or substantively important positive effects. OR both of the following: Two or more studies show statistically significant or substantively important negative effects, at least one study shows statistically significant or substantively important positive effects, AND More studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.

NB: Expert opinion only applicable for 'Minimal Evidence Base' and when 'opinion based on defensible interpretation of theory'.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence. Two or more studies show statistically significant negative effects, at least one of which meets WWC group design standards without reservations, AND No studies show statistically significant or substantively important positive effects.

EXTENT OF EVIDENCE

Extent of evidence (Table IV.4 in Procedures Handbook) is also applied with criteria of:

Medium to large: The domain includes more than one study, AND • The domain includes more than one setting, AND • The domain findings are based on a total sample of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.

Small: The domain includes only one study, OR • The domain includes only one setting, OR • The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

IMPACT (IF SEPARATELY CONSIDERED)

Improvement index:
'The expected change in percentile rank for an average comparison group student if the student had received the intervention'

PROCESS

-

COSTS

Cost data may be included in the intervention report but not shown in the portal

NATIONAL INSITUTE FOR HEALTH AND CARE EXCELENCE (NICE)

Evidence review method

REVIEW APPROACH / EVIDENCE STANDARDS

(Technical method of review + Relevance of method + Relevance of focus)

Systematic reviews of quantitative and qualitative evidence (depending on topics), with statistical meta analysis for quantitative data where possible.

Main focus of the portal's research questions: Issues for the topic area identified in the scope for each guideline.

REVIEW INCLUSION CRITERIA:

Study method: Experimental and quasi experimental studies.

Effect size: Many details in Guidelines Manual.

Other criteria: Inclusion and exclusion criteria explicitly stated in the review protocol, which includes criteria such as:
• PICO
• Study design
• Setting
• Other topic specific exclusion criteria

STUDY IDENTIFICATION:

Systematic search

Evidence claim evidence standards

(Method of review + Included Studies + Nature and extent of totality of evidence)

STRENGTH OF EVIDENCE / BENEFIT

Multiple methods including GRADE rather than a single scale system. The Manual provides an example of a non-GRADE assessment:

• No evidence (not the same as evidence of no effect)
• Weak evidence
• Strong evidence
• Inconsistent evidence.

NATURE AND TOTALITY OF EVIDENCE

Part of GRADE and may be in other evidence summaries.

IMPACT (IF SEPARATELY CONSIDERED)

Part of GRADE and may be in other evidence summaries
Part of QUALY cost calculations

PROCESS

Research evidence beyond effectiveness may be included in the evidence findings.
Studies of uptake of guidance recommendations by practice

Guidance

Creating and communicating practice guidance is the main focus of the work. This involves: the identification of a practice issue by government, consultation on the issues through a public consultation; creation of guidance committee to identify key questions and commission evidence reviews to address them and to then interpret the evidence in terms of other academic and practice and user perspectives and contextual information; public consultation on draft guidance.

The guideline should explain clearly how the Committee moved from the evidence to each recommendation, and should document how any issues influenced the decisionmaking. It should describe the relative value placed on outcomes, benefits and harms, resource use, and the overall quality of the evidence, as well as other considerations of the Committee.

INCLUDED STUDY EVIDENCE STANDARDS

(Technical method + Relevance of method + Relevance of focus)

Every study should be assessed using an appropriate checklist. The quality is then summarised by individual study and, if using the GRADE approach, by outcome across all relevant studies.

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: Guidelines Manual including study appraisal and explanation of various types of synthesis (quantitative and qualitative) depending on the review question.

Manuals on study appraisal: -

Explanation of criteria for appraisal specifically:

COSTS

Cost effective calculations. For health interventions these may be translated in Quality Adjusted Life Years (QALY).

WHAT WORKS CENTRE FOR CRIME REDUCTION: CRIME REDUCTION TOOLKIT

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* **EMMIE framework that measures:** Effect, Mechanism/ theory of change. Moderator variables, Implementation factors, Economic costs. Where there are a number of systematic reviews relevant to a single intervention, EMMIE scores are based on the findings from all of the reviews and the strongest score (for each of the EMMIE elements) is identified and reported as the EMMIE Toolkit Quality/Effect score. Therefore, ratings are not necessarily derived from a single review and might reflect the reported findings from multiple systematic reviews. **Main focus of the portal's research questions:** Approaches. REVIEW INCLUSION CRITERIA **Study method:** Systematic reviews: meta-analysis, narrative review, integrative review, realist review, meta-review. **Other criteria:** On topic and the intervention must have at least 2 included studies and a quantitative outcome measure of crime reduction.	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT Effect Scale: 8 point scale from overall evidence shows an increase to show a decrease in crime. ✓ ✓ Overall, evidence suggests a decrease in crime ✓ ✓ ⚑ Overall, evidence suggests a decrease in crime (but some studies suggest an increase) ✓ Overall, evidence suggests no impact on crime (but some studies suggest a decrease) ✕ ✓ Overall, evidence suggests no impact on crime (but some studies suggest either an increase or a decrease) - No evidence to suggest an impact on crime ✕ Overall, no evidence to suggest an impact on crime (but some studies suggest an increase) ✕ ✕ ⚑ Overall, evidence suggests an increase in crime (but some studies suggest a decrease) ✕ ✕ Overall, evidence suggests an increase in crime IMPACT (IF SEPARATELY CONSIDERED) -	(In development)

STUDY IDENTIFICATION:

Systematic search.

INCLUDED STUDY EVIDENCE STANDARDS

(in this case reviews)
(Technical method + Relevance of method + Relevance of focus)

Five point scale from 'No Information' to 'Very Strong Quality' for each of the EMMIE measures.

Highest Very Strong rating for effect = The review was sufficiently systematic that most forms of bias that could influence the study conclusions can be ruled out.

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: Information about EMMI for systematic reviewers.

Manuals on study appraisal: -

Explanation of criteria for appraisal specifically:

NATURE AND TOTALITY OF EVIDENCE

Number of review studies described.

PROCESS

Mechanism, Moderator and Implementation from systematic reviews rated on 5 point scales from 'No Information' to 'Very Strong Quality'.

COSTS

Very Strong Quality = Marginal or total or opportunity costs (and/or benefits) by bearer (or recipient) estimated.

WHAT WORKS CENTRE FOR LOCAL ECONOMIC GROWTH: EVIDENCE REVIEWS AND TOOLKITS

Evidence review method

REVIEW APPROACH / EVIDENCE STANDARDS

(Technical method of review + Relevance of method + Relevance of focus)

The Centre produces 'Evidence Reviews' and 'Toolkits'. 'Evidence Reviews' have explicit methods sections with a Maryland Scientific Methods Scale (SMS) 3 criteria for evidence. Toolkits sit beneath the 'Evidence Reviews' and are narrative reviews of types of actions (policy design elements) with a SMS 2 criteria for evidence.

Main focus of the portal's research questions: Approaches.

REVIEW INCLUSION CRITERIA

Study method: Primary evaluation studies: Randomized controlled trial, quasi-experimental design, regression discontinuity design, and pre-post designs.

Other criteria: On topic. Other review dependent issues.

STUDY IDENTIFICATION

Systematic search.

Evidence claim evidence standards

(Method of review + Included Studies + Nature and extent of totality of evidence)

STRENGTH OF EVIDENCE / BENEFIT

The Evidence Reviews provide a narrative synthesis on evidence meeting at least SMS 3 level.

The Toolkits provide a narrative synthesis on evidence meeting at least the SMS2 level.

How secure is the evidence?
What does it cost?
How effective is it?

IMPACT (IF SEPARATELY CONSIDERED)

3 point scale on the effects.

NATURE AND TOTALITY OF EVIDENCE

The Evidence Reviews provide a narrative synthesis and report on:

What the Evidence shows.
Where the evidence was inclusive.
Where there was lack of evidence
Lessons.

The Toolkits provide a narrative synthesis on evidence meeting and report on what the policy design element aims to do with a 3 point scale on:

How secure is the evidence?
What does it cost?
How effective is it?

Guidance

Toolkits suggest things to consider if planning to undertake the actions.

It is emphasized that the evidence from these impact evaluations is a complement, not a substitute, for local, on-the-ground practitioner knowledge. The Evidence Reviews outline what tends to work – based on the best available impact evidence – but will not address 'what works where' or 'what will work for a particular individual'. Programmes must be tailored and targeted and an accurate diagnosis of the specific challenges a policy seeks to address is the first step to understanding how the evidence applies in any given situation.

INCLUDED STUDY EVIDENCE STANDARDS

(Technical method + Relevance of method + Relevance of focus)

Five point scale based on Maryland Scientific Methods Scale.
The highest rating of 5 is for randomized controlled trials.
Level 5: Research designs that involve explicit randomisation into treatment and control groups, with Randomised Control Trials (RCTs) providing the definitive example.

Level 4: Quasi-randomness in treatment is exploited, so that it can be credibly held that treatment and control groups differ only in their exposure to the random allocation of treatment.

Level 3: Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual (e.g. difference in difference).

Level 2: Use of adequate control variables and either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group, without an untreated comparison group.

Level 1: Either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group, without an untreated comparison group.

PROCESS

-

COSTS

3 point scale on costs from £ to £££.

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: Guide to Evidence Reviews methods.

Manuals on study appraisal: Guide to scoring evidence using the Maryland Scientific Methods Scale Updated June 2016.

Explanation of criteria for appraisal specifically: -

WHAT WORKS IN REENTRY CLEARING HOUSE

Please note that WWRC pages are being transferred to the National Institute of Justice as part of its CrimeSolutions.gov database. NIJ took the studies in the WWRC and reevaluated them using their rating system, and the ones that met the CS standards can now be found in the Corrections & Reentry section of CS at: https://crimesolutions.gov/TopicDetails.aspx?ID=2

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* Listing of primary evaluation studies and rating of their rigour. **Main focus of the portal's research questions:** Programmes. REVIEW INCLUSION CRITERIA **Study method:** Random assignment or quasi-experimental method with matched groups or statistical controls for differences between groups. Sample size of at least 30 in both the treatment and comparison groups. Study conducted by an independent researcher or published in a peer-reviewed journal. Research studies employing strictly qualitative methods aren't included in What Works. **Other criteria:** On topic STUDY IDENTIFICATION: - INCLUDED STUDY EVIDENCE STANDARDS	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT Studies grouped according to level of Rigor and level of effectiveness for different outcome measures: Strong evidence of effectiveness; Modest evidence of effectiveness; No evidence of an effect; Modest evidence of a harmful effect; and Strong evidence of a harmful effect. IMPACT (IF SEPARATELY CONSIDERED) - NATURE AND TOTALITY OF EVIDENCE List of relevant studies and their ratings for strength of evidence. PROCESS Included where an outcome study also includes process or implementation evaluations to look at how an intervention was implemented and its degree of fidelity to the original program design.	Toolkits suggest things to consider if planning to undertake the actions

(Technical method + Relevance of method + Relevance of focus)

2 point scale of Basic or High Rigor: All studies that meet the methodological requirements for inclusion in What Works are automatically at the Basic level; studies that meet an additional set of requirements are rated as High.

SPECIFICATION OF METHOD

Report specific on portal's studies: -

Manuals on portals evidence base methods: -

Manuals on study appraisal: -

Explanation of criteria for appraisal specifically: Scales only.

COSTS

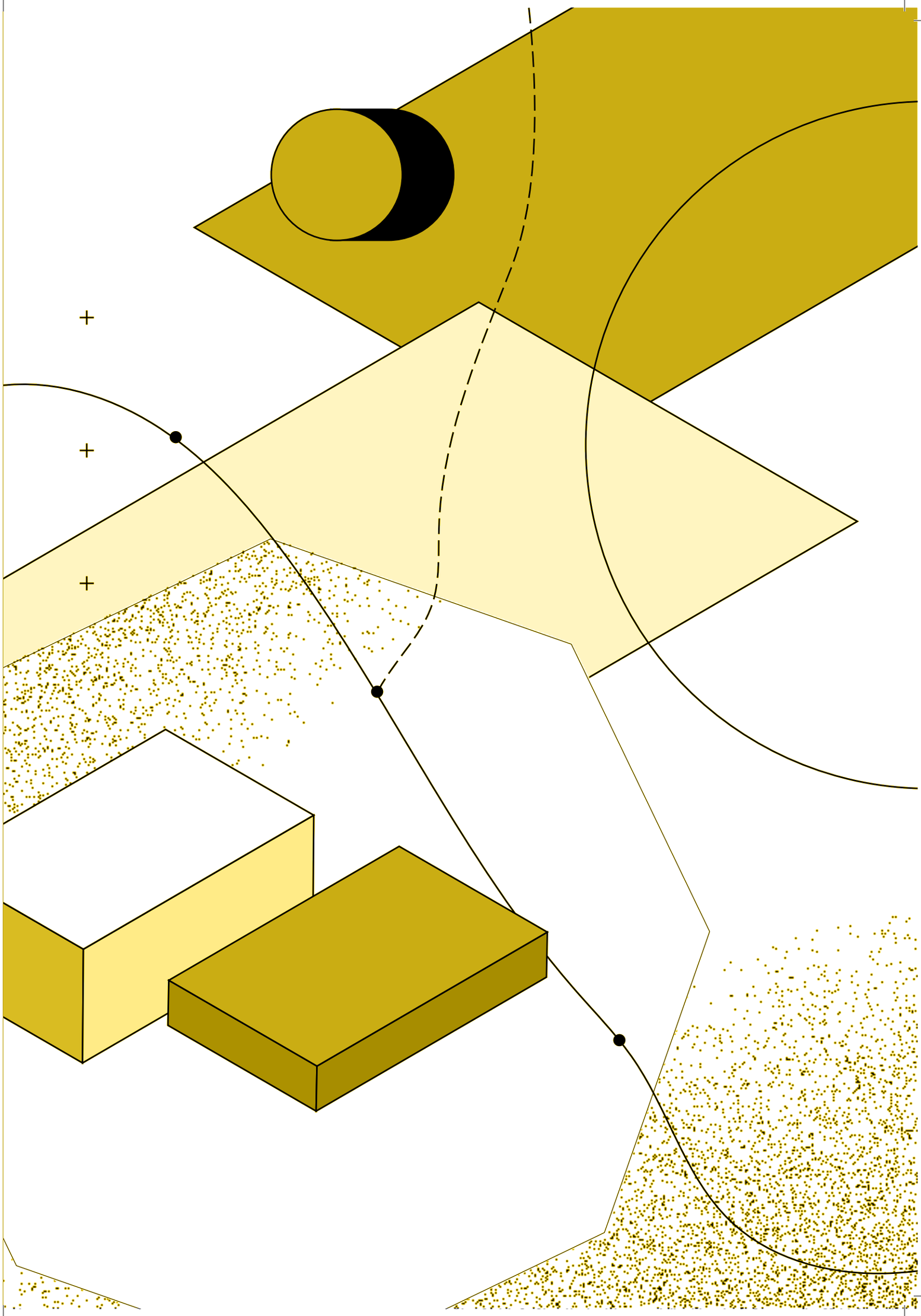
-

WHAT WORKS CENTRE FOR WELLBEING: EVIDENCE COMPARISON TOOL

Evidence review method	Evidence claim evidence standards	Guidance
REVIEW APPROACH / EVIDENCE STANDARDS *(Technical method of review + Relevance of method + Relevance of focus)* Systematic review. **Main focus of the portal's research questions:** Approaches. REVIEW INCLUSION CRITERIA **Study method:** Primary evaluation studies. **Other criteria:** Review dependent. STUDY IDENTIFICATION: Systematic search. INCLUDED STUDY EVIDENCE STANDARDS *(Technical method + Relevance of method + Relevance of focus)* Strong, promising and initial evidence from the primary studies (= high, moderate and low quality evidence / confidence as per GRADE and CERQual guidance). SPECIFICATION OF METHOD **Report specific on portal's studies:** - **Manuals on portals evidence base methods:** A guide to our Evidence Review Method.	*(Method of review + Included Studies + Nature and extent of totality of evidence)* STRENGTH OF EVIDENCE / BENEFIT Strong, promising and initial evidence across the synthesized primary studies (= high, moderate and low quality evidence / confidence as per GRADE and CERQual guidance). IMPACT (IF SEPARATELY CONSIDERED) - NATURE AND TOTALITY OF EVIDENCE Number of studies. PROCESS - COSTS 3 level scale of costs. Cost per person, based on a case study where information is available. Also, used cost per person/ length of impact, e.g.: https://whatworkswellbeing.org/blog/places-spaces-social-connection-and-peoples-wellbeing-what-works/	

Manuals on study appraisal: -

Explanation of criteria for appraisal specifically: -



Centre for Homelessness Impact

Centre for Homelessness Impact
www.homelessnessimpact.org

ISBN: 978-1-9995928-3-7
Crisis
Registered Charity Numbers:
E&W1082947, SC040094.
Company Number: 4024938