



Standardising standards

The case for shared standards in the evidence sector

Jim Vine, Research Fellow, University of Essex

October 2018

There are standards and then there are *standards*. And some of our standards are not standard enough! Ruth Puttick's recent report for the Alliance for Useful Evidence, [*Mapping the Standards of Evidence used in UK social policy*](#), helpfully shone a spotlight on the range of standards of evidence that are currently being used in UK social policy and practice areas. There is clearly an appetite for the use of standardised processes in the evidence ecosystem, and the time is ripe for examining how we collectively approach this standardisation.

This paper provides some background on the different types of standards that can be developed and reviews the benefits that can be obtained from standardisation. I argue that a greater degree of standardisation in social policy and practice evidence — principally by establishing and adopting standards that are common across organisations — would generate even greater benefits than the current organisation-specific standards are able to.

There are many areas where standards could be developed. Two of the key areas are standards for producing different types of evidence and standards for reviewing evidence. As well as outlining several of these, this paper also proposes some other potential areas, such as a management system standard for becoming an evidence-informed organisation and standardised processes for designing and specifying interventions.

This is the sort of proliferation of standards we should welcome and aspire to. For any evidence task that is undertaken at least somewhat frequently by people in different organisations, it would be good to have a single standard available to support the process. Conversely, having multiple standards that seek to achieve roughly the same thing represents duplicated effort with diminishing returns. We want more standards, but not more of the same.

In a collaborative sector, the potential to work together to adopt standardised approaches is relatively high. With few competitive pressures, we can afford to work together to develop standards that can serve the needs of each of our organisations, potentially better even than ones we might make for our own niches. This does not imply any loss of independence or identity. Standards are tools, and shared standards of evidence could comfortably be deployed by different organisations in relation to their respective social policy areas.

What are standards?

At their core, standards are defined ways of doing, thinking about, or managing things (Hogan, Sheehy, and Jayasuriya 2015, p.16). They “*exist principally to provide a reliable basis on which common expectations can be shared regarding specific characteristics of a product, service or process*” (British Standards Institution 2011, para. 4.1.1). They are normally voluntary, rather than mandated by legislation,¹ so are adopted by organisations because they provide some form of benefit, making it easier to do the things they want to do.

One classification that can be informative when considering standards is whether they are internal or external. This is the distinction — typically encountered in business environments — between a company’s own processes that it has standardised, and those things that come from some other standard-setting organisation. Although they are all openly published and available in a way that companies’ internal processes might not be, most (though not all) of the standards of evidence that are documented in *Mapping the Standards of Evidence* could probably be classified as ‘internal’ standards: they are primarily intended for use by the organisations that publish them, whether that is to make funding decisions or to make assessments about the state of evidence for a given intervention.² Terms such as ‘voluntary standards’ and ‘consensus standards’ are sometimes encountered, emphasising important characteristics about the lack of compulsion and communal development processes associated with these types of external standards. (Those rare standards that have the force of government agencies’ regulatory authority behind them can be termed “*mandatory standards*” (David and Greenstein 1990, p.4).)

In his Foreword to *Mapping the Standards of Evidence*, Jonathan Breckon makes a call for us “*to standardise the standards*” (Puttick 2018, p.4). This is a great aim. I read it as a call to move from internal standards, designed primarily for use by one organisation, to external voluntary consensus standards. These standards can be designed through consensus-building with a group of potential user organisations (and open to input from others), and then adopted by all organisations that see value in doing things the standardised way.

Not only is the aim of ‘standardising standards’ a good one, it also seems like it should be an entirely plausible goal considering the context. As Ruth’s findings in *Mapping the Standards of Evidence* highlight, the ‘industry’ in question clearly has an appetite for using standardised approaches, with lots of players having adopted internal standards, with some overlap. They also have the advantage, compared to commercial sectors, that for the most part there is not a great deal of competition between organisations. A rising tide of evidence-based practice is likely to raise all boats, so there is not a strong motivation for any one organisation to seek competitive advantage by having proprietary internal processes that are more efficient than the processes used by other organisations in the sector.

The sector has a lot of potential to move towards the use of external standards, and the benefits of doing so could be substantial.

1. Exceptionally, some standards end up being mandated by legislation or regulations, for example domestic smoke and heat alarms required by the Building Regulations having to conform to particular British Standards (Department for Communities and Local Government 2011). Standards can also become legally relevant in commercial settings in situations where a vendor, for example, makes a claim of a product conforming to a standard; if it turns out to not conform, a purchaser might ultimately be able to refer in legal proceedings to the expectations of the standard that are asserted through such a claim.
2. Having been published there is the possibility of any of these standards being adopted by other organisations wishing to use them. When they are, a case could be made for them potentially serving a hybrid role, combining elements of internal and external standards.

Why we standardise

As primarily voluntary devices, standards are *"tools devised for the convenience of those who wish to use them"* (British Standards Institution 2011, para. 4.1.4). People design them because they believe they will be adopted; and people adopt them because they believe that it will be more convenient to do things the standardised way than any other way. There are several ways in which we might expect a standard to make things more convenient for us, and in doing so *"[promote] productivity and efficiency... [and act] as a catalyst for innovation"* (Hogan, Sheehy, and Jayasuriya 2015).

To consider the benefits of moving towards more standardised standards, it is helpful to reflect on some of the advantages that are commonly identified in relation to standardisation in general. Internal standards are certainly valuable starting points for the sorts of processes that organisations in the evidence sector are following, and deliver some of the benefits of standardisation. For example:

- They **provide consistency (reduce variety)** within their own scope. If different people at Nesta assessed the same grant application against Nesta's *Standards of Evidence* (Puttick and Ludlow 2013), the project should get rated at a consistent level each time. If an evaluation were assessed repeatedly against the EEF's *Padlock Rating System* (Education Endowment Foundation 2016) it should receive the same number of padlocks each time.
- Because they are published, these internal standards support **common expectations** between the organisations that publish them and their stakeholders (e.g. those running the interventions that are assessed using them). If proponents of a particular intervention designed to improve local economic growth submit details of a series of studies relating to the intervention to the What Works Centre for Local Economic Growth, they should be able to form a reasonable idea in advance where on the SMS each study is likely to be rated (What Works Centre for Local Economic Growth 2016).
- They can **reduce the work** of deciding how to undertake a type of task each time. For each intervention that What Works Crime Reduction wants to add to its toolkit, it can deploy the steps detailed in the EMMIE framework, without having to think about how best to rate the intervention (What Works Crime Reduction 2015; Johnson, Tilley, and Bowers 2015).
- Furthermore, the **processes specified can be optimised**, because when you know a process is going to be followed repeatedly it is more likely that you will be able to commit more time to ensuring that all necessary steps are covered, that they are arranged in an efficient fashion, and that unnecessary steps are avoided.

If shared standards were developed and adopted, these advantages would only be likely to be further enhanced and amplified.

- An internal standard creates consistency within an organisation; an external standard across a range of organisations.
- An internal standard supports shared expectations within the organisation, and with its closest partners; an external standard can set the scene for shared expectations that will persist across a broader sector.
- Where internal standards beat case-by-case ad hoc approaches by only needing to be defined once per organisation, external standards can further *"avoid unnecessary duplication"* (British Standards Institution n.d.), and hence achieve even greater economies of scale, by being defined once to solve the same problem for a whole class of organisations undertaking similar tasks.

External standards also have some unique advantages. They enhance interoperability, which is most familiar in the context of technical standards: once a standard has become established, you shouldn't need to worry about whether your Blu-ray player is made by Sony and mine is made by Panasonic; we should be able to borrow each other's discs and play them without problem. The evidence ecosystem for social policy and practice could benefit from a degree of interoperability too. Organisations like the What Works Centres have their particular focuses and expertise, but many social interventions are multifaceted and/or could contribute to outcomes in a range of areas. For organisations delivering interventions there could be lower barriers to recognising and measuring the multiplicity of their impacts if a similar process were in place across social policy domains. Those researchers who conduct studies in more than one sector would be faced with more consistent sets of process and reporting requirements when commissioned by different agencies. This might improve productivity, if researchers' setup effort for each study was lowered due to greater consistency, and for commissioners this could in turn be one example of a mechanism by which external standards might reduce costs. Similar cost savings might be achievable in areas like IT platforms. If multiple organisations were using the same underlying process to assess evidence on interventions, for example, they could also commission a single IT system for handling that work or presenting its findings, which each could deploy for their own sector (no doubt with their own branding for any public-facing elements).

Due to the expectation that they will be deployed in multiple contexts, there is perhaps more pressure for well-designed external standards to be drafted to be clear and unambiguous. Whether internal or external, a lack of ambiguity in a standard helps it to form the basis of specifying to partners what you want them to do. In the case of producing evidence of the effectiveness of an intervention, it could be that you are the organisation delivering the intervention and the partner is an external evaluator. Or alternatively it could be that you are a funder and want to specify that a project receiving funding evaluates to a certain standard. In either case, rather than attempting to detail in full all of the things that this evaluation ought to include, one option would be to start by specifying that the evaluators need to comply with a standard for producing evidence of the effectiveness of interventions (Vine 2016b), and then only needing to provide further details about any specific additions you wish to make.

External standards can also accrue benefits from the range of input into their design. More eyes, and greater diversity in sectoral backgrounds and expertise, are likely to bring more insights to the design of the standard and to reduce oversights and other poor design choices. In the case of the standard for producing evidence of the effectiveness of interventions (Vine 2016b) and associated documents, they were the results of a year's work, including input from a working group with a range of practitioners and a correspondence group with diverse research and evaluation expertise, with dozens of people in total having looked at it.

The breadth of experience and insights that go into creating an external standard can help to *"accelerate innovation across a sector"* (British Standards Institution n.d.). Approaches that have become familiar and demonstrated their benefits in one part of the evidence ecosystem could become part of a standard, which would serve to accelerate their uptake in other areas.

Organisations that use standards can also experience benefits if they want to expand their scope in the future, for example in extending their range of services or entering new markets (Weissinger 2014). An organisation in the evidence ecosystem that has been undertaking assessments of individual evaluations related to its thematic niche, for example, may seek to draw upon standards if it extended into systematic reviews of whole bodies of evidence on interventions, which could give confidence to its funders or other stakeholders. As UK evidence organisations grow, where it would be beneficial to their missions to expand to provide international coverage the use of standards could potentially give confidence to international funders/partners about their ability to consistently deliver outputs with the required qualities.

What we should standardise in the world of evidence

No one standard could reasonably hope to meet all of the needs of the whole evidence ecosystem. Different types of evidence need to be produced; organisations receiving evidence need to assess it; where there are multiple pieces of evidence about a single thing, these need to be combined. Each of these, and more, require distinct processes: a standardised process for producing evidence of the effectiveness of interventions should naturally look different to a standard for assessing how good evidence is on a topic. Rather than aiming for a single unwieldy standard to address all of these legitimate needs, a series of standards could be developed, each with its own scope. Of course, these should be designed to work together well: if you do a decent job of following an evidence-production standard, it would be reasonable to expect that the resultant piece of evidence should tend to be well-received in a complementary evidence-assessment standard.

Two of the main areas where standards could be developed and adopted are around producing evidence and reviewing evidence.

The evidence needs of organisations extend beyond just evidence of the effectiveness of interventions, so evidence-production standards might cover topics such as:

- **Producing evidence of the costs of interventions and quantifying their benefits:** Economic evaluation techniques exist to support decision-making about whether interventions are worth investing in. Even if the intervention has been shown in an effectiveness evaluation to deliver some beneficial outcomes, then these benefits must be weighed against the costs of achieving them, and alternative potential uses for the resources, when considering whether they are a 'good buy'. Some techniques also include converting potentially disparate outcomes into a 'common currency', to assist investment decisions across the breadth of a domain.
- **Producing evidence of implementation feasibility:** For early-stage novel interventions, this type of process evaluation can be helpful to check whether the intervention is deliverable in its current form, and to get a sense of whether any tweaks might make for smoother roll out.
- **Producing evidence of evaluation feasibility:** Another potential type of process evaluation, for investigating in an initial or pilot study whether there are any barriers that will make it difficult to undertake a robust effectiveness evaluation in a main study.
- **Producing evidence of incidence/prevalence:** This would include the case of producing evidence of the number of people experiencing some issue or suffering from some problem that might benefit from interventions. The process would help to establish the level of need (or perhaps in some cases demand) for services/interventions. It might also include the case of establishing the extent of some risk (or this might need a separate process).
- **Producing evidence of opinions:** For those with a focus on interventions, this might include evidence of service users' satisfaction, though the processes could potentially be developed to cover a broader range of opinion-gathering cases.

- **Producing (and using) evidence from models:** Quantitative models, including those that are intended to produce predictive outputs, are increasingly being looked to as sources of evidence in some contexts. Models have features that are distinct from many other types of evidence and require care to be taken in their development and in understanding their limitations.
- **Producing evidence using particular methods:** Method-specific standards could provide details of steps that should be followed when a particular method is being deployed, complementing more general standards on producing evidence. For example, whereas a standard for producing evidence of the effectiveness of interventions can provide an overall process covering a range of possible research methods, a complementary standard could go into details specific to using a randomised controlled trial (RCT) to produce such evidence. This area could also include standards on other types of evidence tools, such as causal chain mapping, and parts of the evidence production process such as statistical analysis plans.

Reviewing evidence could be considered as (potentially) including several elements, which might be helpfully presented as a single standard or split into separate components depending on how organisations feel they might be used in practice:

- **Evidence search processes:** Approaches to identifying and retrieving reports of studies relevant to a subject, and extracting / documenting information from those reports in a structured form. Either as part of this, or as a separate standard, a process for evidence 'gap mapping' could be useful, to establish areas within a field with an absence of evidence.
- **Reviewing studies:** Including forming a view on the robustness of the findings from a given study.
- **Intervention assessment:** Coming to conclusions about an intervention or class of interventions, given various studies that have looked at it. This is distinct from the process of reviewing individual studies, as it would typically involve systematically synthesising evidence from more than one study.

The approaches to reviewing evidence may need to vary dependent upon what type of studies are included. Quantitative findings from multiple studies might be suitable for synthesis using quantitative methods from meta-analysis, for example, whereas qualitative evidence would suit qualitative synthesis techniques.

Other areas related to evidence in which standards might usefully be developed include:

- **Becoming an evidence-informed organisation:** Management system standards exist in other sectors, such as ISO 9001 for quality management and ISO 14001 for environmental management. An evidence-informed management standard could be adopted by practitioner organisations that are committed to making more evidence-informed decisions, adopting practices that have established evidence bases, and possibly contributing to evidence creation where they are considering doing something novel that has not yet accrued evidence.
- **Intervention design and specification:** The standard for producing evidence of the effectiveness of interventions (Vine2016b) adopts a version of the *TIDieR Checklist* (Hoffmann et al. 2014) to ensure that sufficient details are recorded about interventions so that they can be understood (and potentially reimplemented) by evidence users. This sort of information would be valuable to standardise across evidence efforts, such that interventions are thoroughly described in a consistent fashion if they have been through an evidence review process, as well as through evidence creation processes.

- **Outcome measure sets:** Complementing the standards for producing evidence, agreed core sets of outcome measures (with details of how they should be measured, any processing of raw data, etc.) would be helpful in comparing different interventions. They would also facilitate combining results from different studies of the same intervention. Similar sets of outcome measures have been developed for medical trials and are registered under the COMET (Core Outcome Measures in Effectiveness Trials) Initiative (Williamson et al. 2017).
- **Contextual evidence:** Standardising the collection and reporting of information on the context(s) in which a study was conducted or an intervention has been delivered. This would help evidence users to consider whether the results of a given study are likely to replicate in their context and would help evidence synthesisers to investigate variation in effectiveness between contexts.
- **Ethical production (and use?) of evidence:** The ethical conduct of research is clearly an important topic, and there are potentially also ethical issues around how evidence is used. As these are large topics, comprehensive coverage may be more suited to standalone standards than including as elements within (for example) evidence production standards.
- **Sector-specific guidance:** Potential complements to some of the standards, discussing factors relevant to their use in particular sectors, and giving examples from those sectors.
- **Classification systems:** Adopting taxonomies/ontologies of things like different social policy areas or intervention types, perhaps with associated code numbering systems, would allow studies to be coded for easier search and retrieval.
- **Data sharing/analysis code sharing:** With the increasing focus on the reproducibility of studies, sharing data and analysis code is becoming more common. A standard could usefully provide a process to support this being done in the most effective fashion, including features like planning for data release at the design stage, selecting formats that are accessible, and providing sufficient metadata that the data released are meaningful. As well as covering primary evidence production studies, this might also be useful in relation to the data gathered in evidence review processes (in the form of the studies that were identified as part of the literature search, for example).

Clearly this broad range of potential standards would not all be of equal importance to all organisations across the evidence ecosystem. The critical mass of organisations like the What Works Centres, which are already undertaking evidence assessment exercises following internal standards, may make those a particularly fruitful area for early development of consensus-built external standards. If there were appetite for it amongst practitioner organisations, a management system standard to support their journeys to becoming evidence-informed organisations could potentially have great impact, as it is an area where organisations would probably currently struggle to find information on what such an approach might look like. And in terms of standards for producing evidence, economic evaluation would be a particularly valuable complement to producing evidence of the effectiveness of interventions, to provide a framework for considering their costs and benefits.

How to standardise

As mentioned above, standards are primarily voluntary, and that is certainly likely to be the case for any standards of evidence. To create something that is compelling, so that organisations want to sign up to them, the development process is probably as important as the technical content of the eventual document. For the development of the standard for producing evidence of the effectiveness of interventions (Vine 2016b), we followed a process with similarities to BSI's 'PAS' process. The development of the standards was funded by a group of organisations that wanted to sponsor outputs that would ultimately be freely available to any organisation that wished to make use of them.³ The development process was open and sought input from a wide range of stakeholders. It included a working group featuring representatives of the sponsoring organisations amongst others with close interest in their development; a correspondence group representing a broader pool of people, including those with particular technical expertise to contribute; and an open public consultation on a near-final draft to enable anyone to comment before adoption.

As part of the crucial consensus-building, working papers were produced early in the process, covering topics such as the relative strengths of different research methods (Vine 2015a), summarising various sectors' adoption of evidence-based approaches (Howard 2015), and a cross-sector investigation of problems that have been identified in evidence-based practice (and approaches that have been designed to avoid them) (Vine 2015b). These helped to build shared understanding amongst the project's stakeholders, and served to stimulate discussion, building up an understanding of the issues the standard needed to address and features that would help the standard to meet its objectives.

These sorts of standards, intended for adoption by a broad range of organisations, have a particular writing style that is quite distinct from most other reports and documents you are likely to read. A high premium is placed on standards being as unambiguous as possible, and standards have been developed (yes, standards about standards!) that specify how this can be achieved. The style includes, for example, careful and consistent use of 'shall' to indicate requirements and 'should' to indicate recommendations; using the same format for all of the requirements makes standards a bit repetitive, but the lack of sparkle in the prose is traded off for precision and clarity. Features like this contribute to ensuring that standards can support unambiguous assessments of conformity: if there is no room for doubt about whether each point is a requirement or a recommendation, that removes one potential source of disagreement over whether all of the requirements have been followed. Similarly, terminology is usually consistent throughout a standard, avoiding the use of synonyms; this again can make for a drier read, but avoids potential confusion about whether term A and term B are being used interchangeably, or to refer to slightly different things. Sections are numbered and some sections (such as the 'scope') are included in all such standards.

3. This free dissemination is distinct from the typical approach of organisations like BSI and ISO. National and international standards are normally made available on a paid-for basis.

To further avoid potential ambiguity, a finished standard is indicated by a date (year), making it easy for users to reference the version that they are complying with, in case the standard is ever updated. Whilst, as described above, this sort of comprehensive process is likely to lower the rate of errors and omissions, it still does not guarantee that none will remain. Even if the standard is the best it can be at the point of publishing, the world moves on and future updates may be desirable.

The use of different types of standards can be helpful in ensuring a range of user needs are met. The main type, giving the technical details of a process or similar in unambiguous language, is a 'specification'. In the case of the standard for producing evidence of the effectiveness of interventions, the specification was supplemented with a 'guide' style of standard, which mostly provides explanations of why various features of the specification were included and elaboration on points (Vine 2016c). In addition, the definitions of terms were published in a standalone 'vocabulary' standard (Vine 2016a).

One concern noted in *Mapping the Standards of Evidence* regarding the adoption of shared standards is the potential impact on organisations being able to build and maintain their own brand identities (Puttick 2018, p.20). The use of an external standard need not hinder that at all. The Education Endowment Foundation and WWC Crime Reduction (for example) could each have their own toolkits produced using the same standard approach to rating evidence, and they could remain branded as the toolkits of their respective organisations. Organisations that do adopt a standard may wish to acknowledge somewhere that they have done this — e.g., to indicate that the assessments in a toolkit conform to a standard that was used — but such a choice would be optional, and could certainly be subservient to the organisation's own branding. By way of analogy, the carbon monoxide alarm I have in my kitchen has a nice big manufacturer's logo at the top, but also a kitemark and number near the bottom, serving as a claim of conformity to the relevant standard.

Priorities for standardisation

Participation in standardisation processes is voluntary, just as the adoption of standards is. The greater the engagement of evidence organisations, the higher the likelihood of success. The appetite for involvement will naturally depend on the alignment between organisations' interests and the standards that are being worked on. Consequently, this section is inherently quite tentative. It advances an argument for a particular set of priorities, but does so in the knowledge that if collectively the organisations in the evidence ecosystem would favour a different sequencing, it would be better to move forward in line with their preferences in the initial stages of establishing external standards. We may need to seek consensus even in the choice between potential consensus standards.

The lion's share of the 18 standards in *Mapping the Standards of Evidence* provide approaches for assessing evidence in one way or another. This is clearly a central focus of many of the What Works Centres, amongst other organisations that look to provide sectors with information on the current state of evidence in their niches. Standardisation in this area would represent the single biggest opportunity to bring organisations together to build consensus in an area that is of current relevance and to standardise a substantial area of current practice. The flip side of that is the existing processes that are embedded in organisations and embodied in the wealth of evidence assessments undertaken to date, potentially introducing inertia and disincentivising change.

If possible, that inertia should be resisted. With evidence review activity ongoing, the weight of assessments conducted under internal standards will only increase with time, until such time as an external standard is adopted. The point at which a new approach is adopted, developed as a shared standard through a consensus process, need not imply that all that preceded it should be instantly discarded. In practical terms, organisations could continue to present the findings of their 'legacy assessments' alongside those undertaken using their newly adopted method, and each organisation could make its own decisions about how high a priority it places on updating those assessments, not least because the effort involved in those updates would vary depending on the degree of similarity between the external standard and an organisation's previous internal one. For some organisations that prioritisation decision could even result in letting legacy assessments stand indefinitely, perhaps only being updated when new evidence on a topic is published.

At the same time, it would be nice for consensus-building approaches to be deployed to start addressing questions that few or no organisations currently have established answers to. This would help to set the scene for a future in which the evidence ecosystem works collaboratively to innovate through the mechanism of standards and to test how that process works. One option would be to start by identifying two or more areas for standardisation, with different characteristics, such as different complexities. Various types of evidence production, for example, are relatively well-known quantities, and whilst producing a standard for producing a type of evidence is not trivial, the complexity would be significantly moderated by existing examples of how studies of that type have been conducted. Conversely, the higher novelty factor in something like an evidence management system would test other parts of the standard-setting process, as well as having potentially higher payoff in proportion to the extent of any current demand for such a process being relatively unmet.

Conclusions

The UK evidence sector as we would currently recognise it is still relatively young. The Standards of Evidence timeline in *Mapping the Standards of Evidence* goes back to 2000 (and is sparsely populated before 2011) (Puttick 2018, p.8), and the What Works Network was only established in 2013 (Gold and The What Works Team 2018). And yet this sector has already shown a great appetite for developing standardised processes, recognising that these are essential for consistent delivery of programmes of work, with an associated proliferation of standards.

It is now time for standards of evidence to proliferate in a different direction and using different means. There is a wide variety of topics where the evidence ecosystem could benefit from having a standardised process to support its practices. Different types of evidence are required, and a standard for undertaking economic evaluations would contain different elements to one for producing evidence on opinions, for example. The evidence reviewing process can be thought of as having distinct components, and these could be standardised either in whole or as a series of complementary parts. And many more areas, from intervention design to management of organisations that want to improve their use of evidence, could all fit within an expanded horizon of standards of evidence.

These new standards of evidence should be built through consensus processes, bringing together organisations from across the sector and being open to even broader input into their development. Their adoption will inevitably be voluntary, so diverse participation will be crucial, not just for ensuring that the standards are of high quality and meet organisational needs, but also to build a sense of buy-in from the outset of the process.

One possible initial programme of activity for developing shared voluntary consensus standards of evidence might look like the following, commenced in parallel:

- 1. Standards for Reviewing Evidence.** A sequential approach could be useful, starting with processes for identifying, retrieving, extracting and documenting evidence, before moving onto the actual approaches to establishing an assessment of evidence from a single study, and combining the results from multiple studies.
- 2. Standards for Producing Evidence.** Evidence organisations should select a type of evidence for which there is a reasonable degree of shared interest. Given many evidence organisations' interests in informing decisions of whether interventions should receive investment, economic evaluations may be a strong contender.
- 3. Evidence-Informed Management Standard.** This, or another 'blue skies' standard, could complement the others, and would help to build understanding of the potential for standardisation to support innovation in the evidence sector.

If these are steps 1, 2 and 3, we can perhaps add a zeroth step, to help encourage future standards to be made as community efforts wherever possible:

- 0. Moratorium on new internal standards.** Organisations in the evidence ecosystem should commit to refrain from developing internal standards, except where it is very clear that they relate to things that are entirely specific to that organisation's context.

The evidence ecosystem has opportunities to build on its existing set of internal standards and to strike out in new directions. Each standard developed would broaden the areas of the evidence system where people and organisations could rely on a shared approach in support of their vital work. Evidence organisations should do this with a sense of community, drawing on the strengths of a sector that is largely insulated from the need for organisations to compete with each other, and equipping the world with the best set of tools to ultimately bring high quality evidence to bear across the breadth of social policy areas.

References

- British Standards Institution (2011) 'BS 0:2011 - A Standard for Standards - Principles of Standardization.' London: British Standards Institution. <http://www.bsigroup.com/LocalFiles/en-GB/standards/bs0-pas0/BSI-BS0-Standard-for-Standards-UK-EN.pdf>
- British Standards Institution. n.d. 'PAS: The Sponsored Approach for Developing Consensus Based Standardization with the UK's National Standards Body.' Accessed July 28, 2015. http://www.bsigroup.com/LocalFiles/en-GB/standards/BSI_PAS_6pp_DL_AW_LR.PDF
- David, P. A. and Greenstein, S. (1990) The Economics Of Compatibility Standards: An Introduction To Recent Research. 'Economics of Innovation and New Technology.' 1 (1-2): 3-41. <https://doi.org/10.1080/10438599000000002>
- Department for Communities and Local Government (2011) 'The Building Regulations 2010: Approved Document B: Fire Safety, Vol. 1: Dwelling Houses.' NBS.
- Education Endowment Foundation (2016) 'Classification of the Security of Findings from EEF Evaluations.' https://educationendowmentfoundation.org.uk/public/files/Evaluation/Carrying_out_a_Peer_Review/2016_Classifying_the_security_of_EEF_findings.pdf
- Gold, J. and The What Works Team (2018) 'The What Works Network - Five Years On.' What Works Network. <https://www.gov.uk/government/publications/the-what-works-network-five-years-on>
- Hoffmann, T. C., Glasziou, P.P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D.G., et al. (2014) Better Reporting of Interventions: Template for Intervention Description and Replication (TIDieR) Checklist and Guide. 'BMJ.' 348 (March): g1687. <https://doi.org/10.1136/bmj.g1687>
- Hogan, O., Sheehy, C. and Jayasuriya, R. (2015) 'The Economic Contribution of Standards to the UK Economy.' London: BSI. <https://www.bsigroup.com/LocalFiles/en-GB/standards/BSI-The-Economic-Contribution-of-Standards-to-the-UK-Economy-UK-EN.pdf>
- Howard, W. (2015) 'State of Evidence-Based Practice in Other Sectors.' Working Paper WP2. London: HACT. <http://www.hact.org.uk/sites/default/files/WP2%20-%20State%20of%20evidence-based%20practice%20in%20other%20sectors%20v1.1.pdf>
- Johnson, S. D., Tilley, N. and Bowers, K.J. (2015) Introducing EMMIE: An Evidence Rating Scale to Encourage Mixed-Method Crime Prevention Synthesis Reviews. 'Journal of Experimental Criminology.' 11 (3): 459-73. <https://doi.org/10.1007/s11292-015-9238-7>
- Puttick, R. (2018) 'Mapping the Standards of Evidence Used in UK Social Policy.' London: Nesta. <https://www.nesta.org.uk/report/mapping-standards-evidence-used-uk-social-policy/>
- Puttick, R. and Ludlow, J. (2013) 'Nesta Standards of Evidence.' London: Nesta. http://www.nesta.org.uk/sites/default/files/standards_of_evidence.pdf
- Vine, J. (2015a) 'Introduction to Research Methods and Considerations of Their Relative Strengths.' Working Paper WP1. London: HACT. <http://www.hact.org.uk/sites/default/files/Intro%20to%20research%20methods%20-%20Stds%20of%20Evidence%20WP1%20v1.1.pdf>
- Vine, J. (2015b) 'Learning Lessons from Challenges Encountered in Evidence-Based Practice in Other Sectors.' Working Paper WP3. London: HACT. <http://www.hact.org.uk/sites/default/files/WP3%20-%20Lessons%20from%20challenges%20other%20sectors%20v1.0.pdf>
- Vine, J. (2016a) 'General Requirements for Evidence – Part 1: Vocabulary.' Standard StEv1-1:2016. Standards of Evidence. London: HACT. [http://www.hact.org.uk/sites/default/files/StEv1-1-\(2016\)%20General-Vocabulary.pdf](http://www.hact.org.uk/sites/default/files/StEv1-1-(2016)%20General-Vocabulary.pdf)
- Vine, J. (2016b) 'Standard for Producing Evidence – Effectiveness of Interventions – Part 1: Specification.' Standard StEv2-1:2016. Standards of Evidence. London: HACT. <http://www.hact.org.uk/sites/default/files/StEv2-1-2016%20Effectiveness-Specification.pdf>
- Vine, J. (2016c) 'Standard for Producing Evidence – Effectiveness of Interventions – Part 2: Explanation and Elaboration.' Standard StEv2-2:2016. Standards of Evidence. London: HACT. <http://www.hact.org.uk/sites/default/files/StEv2-2-2016%20Effectiveness-Explanation.pdf>
- Weissinger, R. (2014) Economic Benefits of Standards. 'PIK - Praxis Der Informationsverarbeitung Und Kommunikation.' 37 (3): 183-188. <https://doi.org/10.1515/pik-2014-0015>
- What Works Centre for Local Economic Growth (2016) 'Guide to Scoring Evidence Using the Maryland Scientific Methods Scale.' What Works Centre for Local Economic Growth. http://www.whatworksgrowth.org/public/files/Methodology/16-06-28_Scoring_Guide.pdf
- What Works Crime Reduction (2015) 'What Works in Crime Reduction Toolkit: Method Statement.' http://whatworks.college.police.uk/toolkit/Documents/CLEAN_method_statement_1115.pdf
- Williamson, P. R., Altman, D.G., Bagley, H., Barnes, K.L., Blazeby, J.M., Brookes, S.T., Clarke, M., et al. (2017) The COMET Handbook: Version 1.0. 'Trials.' 18 (3): 280. <https://doi.org/10.1186/s13063-017-1978-4>

Acknowledgements

My work on this paper was partially supported by grant number ES/L011859/1, funded by the Economic and Social Research Council, establishing the Business and Local Government Data Research Centre and the Catalyst Project, initially funded by HEFCE and now administered by the Office for Students. I am grateful to **Ruth Puttick** for her thoughtful comments on an earlier version of this paper, and indeed for producing the *Mapping the Standards of Evidence Used in UK Social Policy* report that so clearly illuminated the proliferation of standards in use across the evidence ecosystem. Any errors that remain are, of course, the responsibility of the author.

About the author

Jim Vine's research experience spans a range of local and national governmental contexts, as well as charitable and business sectors. He has undertaken research relating to both policy and practice, working in research charities for a decade before taking up a research post in academia. Whilst working at HACT he developed a set of Standards of Evidence for producing evidence of the effectiveness of interventions, funded by Public Health England and a group of housing associations – Bromford, Look Ahead, Metropolitan, Sanctuary Supported Living and Trafford Housing Trust.

jim.vine@essex.ac.uk

[@jim_vine](https://twitter.com/jim_vine)

If you'd like this publication in an alternative format such as Braille, large print or audio, please contact us at: information@nesta.org.uk



Funded by:

nesta



58 Victoria Embankment
London EC4Y 0DS

+44 (0)20 7438 2500

www.alliance4usefulevidence.org
www.nesta.org.uk

Nesta is a registered charity in England and Wales with company number 7706036 and charity number 1144091.
Registered as a charity in Scotland number SCO42833. Registered office: 58 Victoria Embankment, London, EC4Y 0DS.

