TReQ Tools: How to Improve Transparency, Reproducibility and Quality in Energy Research (version 0.1)

Gesche Huebner[1] and Michael Fell[2], University College London

Version 0.1 published 20 January 2020

This working paper provides a toolkit to help energy researchers conduct transparent, reproducible (where appropriate) and quality research -- or TReQ research for short.

Energy use is key to global challenges such as climate change, while also playing an important role in the day-to-day lives of people and communities. But how sure can we be that the research done in these areas is providing reliable findings?

Unlike in fields such as medicine, energy research has not adopted many of the tools and practices needed to maximise the transparency, quality and reproducibility[3] of studies. The consequence of this is that evidence-based policy and practice may be built on shaky foundations.

This document aims to set out the principles and tools that all energy researchers should now consider bringing to their work. It provides practical instruction on approaches such as pre-analysis plans, reporting guidelines and systematic evidence reviews that are widely used in other fields to support TReQ research.

Crucially, it draws on the recognition that real-world research can be messy and unpredictable – and therefore considers how best to design and conduct resilient research under these circumstances. Applicable for quantitative and qualitative approaches in a wide range of disciplines, we hope this developing document will be a useful reference for energy and buildings researchers everywhere, whether in training or in practice.

Would you like to get involved? **We are keen to develop this resource collaboratively with others in the energy research community. If you would like to get involved, either as a reviewer or to contribute new material, please contact us on g.huebner@ucl.ac.uk and michael.fell@ucl.ac.uk.**

[1] g.huebner@ucl.ac.uk
[2] michael.fell@ucl.ac.uk
[3] For this resource, we have adopted the following definitions (based on https://www.nature.com/news/1.19970): A study is considered reproduced when its findings are confirmed when collecting new data using highly similar methods and materials. A study is replicated when it is repeated exactly, i.e. using the same data, code, etc.

Contents

Versions of this working paper

We expect to add to and refine this working paper over time. Please visit our OSF page for a link to the latest release, as well as a link to a Google Doc of the version currently under development, which you can comment on.

Specific challenges in the energy research field

Energy research is very multidisciplinary: we find engineers, architects, mathematicians, economists, sociologists, anthropologists, statisticians, and others all researching in this field. They use a multitude of methods, such as interviews, focus groups, surveys, field and lab experiments, case studies, monitoring and modelling.

It is a crucial research field given the imminent need of reducing our carbon emissions. Policy needs high-quality, reliable evidence, including reviews of evidence. Given this, along with constraints on available research funding, one might argue that we should deliver the most robust, transparent, and reproducible research and synthesize it in a systematic way. However, we are not as far advanced as other, more homogenous disciplines when it comes to reproducibility. Why?

The diversity of methods and research designs that we use likely plays a role – for many of these, there isn't a clear gold standard as, for example, exists for experimental research. Also, so far we have been spared the feeling of a crisis; there has been no big failure of reproducing an existing, established finding; no one has been exposed as a fraud, etc.

In addition, many of us work with commercially or personally sensitive data that are difficult to share. For example, smart meter data are personal data and making the data freely publicly available is tricky given the legislation around it.

Also, the energy sector is changing rapidly. Other disciplines are interested in establishing general principles, for example, research on perception within psychology, whereas research in the energy area is often much more focused on the current situation, fully aware that ten years down the line, things will have changed. In fact, much of the research is aimed at helping deliver change to meet climate change targets. For a lot of our research, contextual factors play a much larger role than fundamental factors – and so when the context changes, the findings will change. Also, the context in different countries is quite different and so results from another country might not be transferrable irrespective of how well the study was conducted.

Another factor that makes it harder to reproduce previous research is that (especially) field trials are extremely time- and money-consuming, more than in many other disciplines. For example, when wanting to test if specific energy tariffs save social housing tenants money, it is a different magnitude of complexity: there is the need to get a housing provider to agree to run a study; an energy supplier needs to come on board; fieldworkers need to be recruited and trained, etc. We are speaking about a couple of million pounds just for the fieldwork. Similarly, to build an energy system model of the UK, such as the UK Times Model, easily takes several years. This complexity doesn't mean that we are excused from reproducibility but might explain why we have made fewer attempts.

Finally, there are a number of research areas where reproducibility, in a strict sense, might not even be an appropriate term to discuss. Many qualitative and participatory research projects focus on specific case studies where with different participants and different researchers you would expect to find different results.

So far, so tricky. But despite these difficulties, we argue that energy researchers can -- and should -- be doing much more to integrate TReQ research best practice into their work. Luckily, there is a set of simple tools which we think almost all energy researchers should now at least consider employing. The rest of this document is a brief introduction to what these tools are, why and how they are used, and signposting to further information.

TReQ Tools

For each tool, this section explains why, how and when it is used, benefits and drawbacks for researchers, and provides links to further resources.

Preregistration of research protocols and analysis plans

- Involves setting out in advance how you plan to carry out and analyse research, and (when applicable) what you expect to find
- Adds credibility to your results because you can't be accused of changing your analyses/expectations afterwards to fit the data.
- Also helps with planning and getting early input of research design and analysis.

Why?

A shadow looms over research -- the problem of under-reporting of so-called 'null' results. Let's say a researcher has a reasonable belief that a certain intervention will have a certain effect. They test it out, and don't find the effect. Ideally, they would write up that research and it would be published in the same way as any other work -- after all, other researchers should know that the intervention has been tried, so they can either try a different intervention or try testing it in a different way.

However, there is a perception amongst some editors, reviewers, and researchers that publications of the kind described above are less valuable -- that they add less to our knowledge than those which 'work out'. They may be less willing to accept, or recommend acceptance, of such papers. In anticipation of this, or because they themselves undervalue the contribution of such reports, researchers may also be less likely to write papers reporting 'null' results. The upshot of this is the well-known 'file drawer problem' or publication bias, whereby there is under-reporting of findings that do not meet common standards of statistical significance (e.g. $p<0.05$) (see the Wikipedia page for more).

There exists, therefore, a motivation to look for 'positive' findings in research in the expectation (justified or not) that this will make for more publishable and impactful papers. For example, a researcher might repeat an analysis with many different combinations of variables in the hope that some show certain desired statistically significant results (e.g. that certain variables are related). This is known as p-hacking. There may also be a temptation to present certain results as having been expected all along -- so-called HARKing (or Hypothesizing After Results Known). The trouble is that, unless these actions are all fully disclosed, they can lead to findings being presented with an impression of greater confidence than is in fact warranted.

Preregistration of a 'pre-analysis plan' (PAP) is one way of mitigating against these problems. This outlines what the key outcome measures are and a step-by-step overview of how the analysis will be conducted. For example, in some quantitative analyses, this might include what variables will be included in what statistical tests, how missing data and outliers will be handled, and so on. This can then be put online, with a certified time-stamp and registration number. They can either be immediately shared publicly, or kept private until some later date (such as publication of a paper).

If applicable to the research approach, PAPs also give us the opportunity to pre-register our hypotheses -- essentially to say in advance what results we are expecting (often based on theory or previous evidence). This is important because we can then clearly demonstrate, if we find a certain result, that we always had this expectation and didn't just 'retrofit' it after the fact.

Benefits for researchers:

- Get early input
- Think through analysis and avoid missing important data
- Signal you are doing work in this area
- Streamlines the actual data analysis (as you have a plan)

Why not?

In some fields, such as medicine (where PAPs originated) and even psychology and economics, there is considerable agreement over what the key outcome measures are, how they can be measured, and the kinds of analysis used to detect effects. Across energy research there tends to be less consensus around such matters, so it is less straightforward to know exactly what a PAP should include. It may even be hard to know exactly what outcomes measures you will have access to until the dataset arrives. However, these tricky decisions will always have to be made at some point, and putting them down in a PAP can make it easier to spot early on if certain key measures could be missed, that can then be addressed sooner rather than later.

There is also an argument that the time required to write a PAP takes away from time that could be spent on more valuable research activities. However, it is our experience that PAPs are actually timesavers in the long run. Instead of testing this or that approach to cleaning and/or analysis when the data arrive, the best approach has already been decided upon and set out in advance. While of course deciding on this approach takes time, it is quicker to do this when the data are not there to act as a distraction. And when the data are available, all you have to do is follow the guide you have already written for yourself.

Another concern is that employing a PAP will stifle the ability to spot unexpected associations, or deal with data which is somehow different in form to what was expected. The key point to note here is that, while it is important to report on the

analyses set out in the PAP, the analysis does not have to stop there. Doing further or different analysis is completely fine -- it is just important to be clear where it differs from what was specified and, ideally, to make it clear what this means for the interpretation of the results.

For example, imagine that a researcher planned to do a statistical comparison of two conditions, but finds that they don't have enough cases in one of them to do a meaningful statistical analysis. Instead, they may decide to do a visual inspection of the data. Importantly, they are completely free to do this -- but just have to explain why (and they can't claim this was what they planned all along).

A final concern (for now) is that, especially for new researchers, it is important to 'get your hands dirty' with the data. It may be hard for someone with no experience of statistics to set out a useful PAP without having had the opportunity to try things out first. We suggest the best response to this, if possible, is to use an alternative existing (or synthetic, realistically made up) dataset for this kind of exploration, and do your best to write a PAP for any subsequent work you intend to write up. As above, you are still free to vary it if you like, but you have the added strength of a PAP 'in the bank' to increase the credibility of your findings.

For what type of research?

The use of PAPs is most important in deductive (theory-testing) quantitative research, but the principles are applicable to any form of research, including qualitative and modelling approaches. For example, a researcher may wish to set out in advance what sort of codes they will use in qualitative analysis, along with material such as interview topic guides. Systematic reviews employ a similar approach through the publication of a [review protocol].

When in the research process?

At the latest, pre-registration should be done before you start analysing the data -- and really before the data are even received. However, best practice is to be working on your PAP during the course of your research design, and certainly before data collection. This is because writing the PAP may suggest important data you will need in your analysis that you might not otherwise have planned to collect. You will also have more time to get input from others (e.g. expert advisors) on your planned analysis.

How?

You will first need to decide where you want to register your PAP. We recommend AsPredicted.org -- this website allows you to fill out a simple form, add co-authors, and save in a time-stamped version that you can either make publicly available immediately or at some later stage (e.g. on publication of a paper). You can see an example of a completed AsPredicted registration here. However, other research approaches may be better suited to different PAP templates and services. The Open

Science Framework website maintains a useful list of templates, including those specific to qualitative research and secondary data analysis.

When you come to write up your research, it is important to provide details of your PAP (e.g. hyperlink and registration number). You should also make it clear if you have diverged from it at any point, and why.

Further reading

A Guide to Pre-Analysis Plans by Erica Chuang and Stephanie Wykstra provides information on how to put together a more detailed PAP.

How good are pre-analysis plans in practice? and lessons for writing/reviewing your next one by David McKenizie on World Bank Blogs.

Are consumers willing to switch to smart time of use electricity tariffs? The importance of loss-aversion and electric vehicle ownership is a paper by Moira Nicolson, Gesche Huebner and David Shipworth which employs a PAP.

Attractiveness of a broader range of TVT tariffs and their marketing is another example of a PAP registration.

Open data and open code

- Allows identification of errors (intentional and unintentional)
- Increases visibility of the research
- Avoids duplication of data collection
- Allows reusing of data for research, education, and training

Why?

Making data and the corresponding software code publicly available is a key aspect in preventing data fraud and in promoting scientific integrity and replication studies. Errors (intentional or unintentional) in analysis could be identified by others. The real strength of publication of datasets and computer code is in combination with PAPs: If neither data collection tools, variables to be used in analyses nor exclusion criteria have been pre-specified, uploaded data could still be manipulated. Together with PAPs, though, cherry-picking of results and other forms of p-hacking could be uncovered.

There are further benefits to depositing data: e.g. it increases the visibility of the research as data sources can be cited; there are a number of repositories where data sources are issues with a DOI (digital object identifier) and journals such as Nature Scientific Data that have so-called 'data descriptors' as their main publication type. Making data and code available is economic as they be re-used, it can give further discoveries that the original authors had not thought of, and it can help to establish collaboration. The FAIR principles state that data should be Findable, Accessible, Interoperable, and Reusable:

- **Findable.** Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.
- **Accessible.** Once the user finds the required dataset, they must be able to actually access it, possibly including authentication and authorisation.
- **Interoperable.** The data usually need to be integrated with other data. In addition, the data need to interoperate with as wide as possible a variety of applications or workflows for analysis, storage, and processing.
- **Reusable.** To be able to reuse data, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

Why not?

A particular challenge for energy research is that projects are often done with a commercial partner, e.g. a utility company who for competitive reasons might oppose publication of data sets. In the ideal case, this would be addressed prior to

commencing any research and an agreement reached on which meta-data (if not the total data) could be shared.

Data need to be anonymized prior to publication which costs time and / or resources. However, many data sets are already collected in an anonymous way (e.g. survey data using an online platform), or need to be anonymized to comply with GDPR rules. For example, smart meter data are personal data and hence difficult to make publicly available. A current EPSRC funded project (SERL) {LINK} manages to get around this issue by granting access to data in a (currently under development) secure virtual environment. Other data, such as video recordings, can also not be shared – however, a workaround would be to publish a transcript and to erase anything that might identify the person.

Some energy demand studies result in very large data sets. While many repositories are free to use, there is sometimes an excess charge for very large data sets. E.g. Dryad charges $50 for each 10 GB beyond the initial 20 GB (in 2019). However, in the light of what research actually costs, the deposition cost is negligible. The Open Science Framework publishes data sets up to 5 GB for free.

The drive to publish in competitive academic environments might also play a role: Data collection can be a long and expensive process and researchers might fear premature data sharing may deprive them from the rewards of their effort including scientific prestige and publication opportunities.

Open code might not be considered helpful if the software itself needs to be paid for. However, whilst this will limit access for those outside of research / academic institutions, those based in these institutions will most often have access to the software.

For what type of research?

Sharing data and code is most common in quantitative research, but would be applicable to any type of research. For example, codes and codesets used as part of thematic analysis in qualitative research can be shared.

Sharing code can be understood in a wide sense – i.e. any instructions on how to analyse the data in question.

When in the research process?

After you have received the data; for code – sharing can occur earlier if for example, using a mock data set to set up the code. Github makes it easy to keep track of various iterations of code.

How?

1. Identifying a suitable data repository.

The webpage re3data.orgprovides a list of more than 2000 research data repositories that cover research data repositories from different academic disciplines; however, this includes institutional webpages which cannot be accessed by everyone. Nature Scientific Data publishes a list of recommended data repositories on their webpage . Whilst there are no recommended data repositories for energy, repositories for social sciences, related disciplines such as physics, and general use are listed.

Similarly, a range of platforms exist for hosting and sharing code, the most well known one probably github.

Some repositories are suitable for a range of uses under one project, such as the Open Science Framework which allows preregistration of studies, depositing of data and code.

2. Anonymizing data

It needs to be ensured that confidentiality of research participants is safeguarded, i.e. all data need to be de-identified. Name, date of birth, and address would always need to be deleted or at least to be reduced in precision, e.g. reducing date of birth to year only. The UK Data Archive gives a good overview on what to consider in anonymization (https://www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation.aspx).

3. Preparing a meta document / code documentation

A meta-document or 'ReadMe' document needs to be prepared to help others to understand which variables measured, how and when. For code, a meta-document might be advisable, but as the minimum, the code needs to be sufficient commented.

4. Uploading the data / code

You might prefer to upload multiple versions of data, e.g. the raw dataset and a processed one. We would advise to always upload a raw data version where you haven't done any changes (apart from removing identifiable features). Make sure to follow the FAIR principles.

Further reading

The FAIR Guiding Principles for scientific data management and stewardship by Wilkinson et al.

The importance of open data and software: Is energy research lagging behind? by Pfenninger et al.

Recommendations for the Role of Publishers in Access to Data by Lin and Strasser.

Thinking about the Coding Process in Qualitative Data Analysis by Elliott -- indicates the sort of details from qualitative work that could be published.

Assessment of the Space Heating and Domestic Hot Water Market in Europe—Open Data and Results by Pezzutto et al. is an example of open data sharing.

An example of open data and code from the Renewable Energy and the Smart Grid (GREEN Grid) project.

Preprints of papers

- A preprint is a full draft of a research paper that is shared publicly before it has been peer reviewed
- It allows receiving feedback before submission to a journal and hence improving the paper prior to 'official' review processes
- Allows you to show earlier that you are working in a certain area

Why?

Preprints have the potential to bring substantial benefits to the individual researcher and the research community.

Many preprints are issued with a DOI, making them a citable document. By posting a preprint with your research results, you can firmly stake a claim to the work you've done at an earlier time than a regular journal article with its long lead times to publication. This not only allows you to show that you have done certain work at certain times, but you might also be able to use preprints in grant applications to point to finished but not yet published work. It increases your visibility in the research community, and can help to initiate collaborations.

Preprints allow to get early feedback on a full manuscript. Other researchers can potentially point out errors or flaws in the logic of the paper, suggest additional studies or analyses, or point you to relevant work you might have missed.

Usually, a preprint is followed by a full research paper. You can – and actually should – link your preprint to the final published article meaning that your preprint can contribute to bringing new readers to your eventually published paper.

These individual benefits all help the research community as a whole. The early visibility means that others working on a similar topic will not do the exactly same study but can build on your findings, i.e. making research more economic (of course, they might also want to exactly run the same study to see if they can replicate findings, again allowing this earlier than if having to wait for a full journal publication. The increased scrutiny from potentially many more people than the usual 2, 3, 4 journal reviewers can improve the study and hence contribute to ensuring science progresses on the basis of valid, reliable research. Preprints can help to build collaborations. And finally, preprints are per default open access, i.e. everyone can access them.

Why not?

One common concern is that by putting your study early out to the world, someone might repeat the study and beat you to publication. However, this does not outweigh the benefits of preprints and constitutes a small risk. To have the necessary resources at the moment in time to repeat the study is not that likely; maybe even

less so in the energy field where many projects involve external partners such as councils, housing association, etc. Also, because your preprint is posted publicly with a DOI, it becomes a permanent academic record that can be cited and indicates clearly who first did a specific study.

It used to be the case that publishing a preprint might mean that a journal would not consider it as 'unpublished', a common condition for considering a manuscript. However, most journals now allow preprints (see this Wikipedia entry on journal policies regarding preprints).) Nature Research states explicitly on its webpage "Nature Research journals encourage posting of preprints of primary research manuscripts… Posting of preprints is not considered prior publication and will not jeopardize consideration at Nature Research journals."

For what type of research?

Preprints are applicable to any kind of research where the output is a manuscript for publication.

When in the research process?

At the end – i.e. after having run the study, analysed the data, and written the manuscript.

How?

A range of sites exists that hosts preprints. Probably the oldest and one of the most well-known is arXiv.org, a preprint server for the physical sciences. Other domain specific exist such medRxiv (medical studies), bioRxiv (life sciences), chemRxiv (chemistry), and Humanities Commons (humanities). The Center for Open Science (as OSF Preprints) hosts preprints for various disciplines and allows searching for preprints from a variety of other preprint providers such as arXiv, bioRxiv, PeerJ, CogPrints and others. More than two million preprints are indexed and can be searched by author or keywords, or filtered by subject, preprint server etc. Uploading a preprint is straightforward and upon publication, it is issued with a DOI. Following acceptance in a journal, that DOI is added. Preprints cannot be deleted but they can be changed.

Further reading

Preprints for the life sciences by Berg et al. discusses the use of preprints in another area where they have been less widely used.

Smart Local Energy Systems (SLES): A conceptual review and exploration by Ford et al. is an example of a preprint posted on SocArXiv.

Reporting guidelines

- Widely agreed checklists of what details you should be reporting for different kinds of research studies.
- Help give you confidence that you are reporting all the necessary details for your findings to be understandable and reproducible.
- Make it easier to draw on your findings in evidence reviews.

Why?

When you are writing a report of your research, it can be difficult to decide which details are important to include, and which you can afford to leave out. There may even be a risk that you forget to include some important details of the work you conducted. This poses a problem for a number of reasons. Without all the necessary details, it can be hard for readers to judge how well-founded your conclusions are. It may make your work difficult or impossible for others to reproduce. And future evidence reviews may have a hard time integrating your method and findings.

To help address these problems, sets of reporting guidelines have been developed spelling out precisely which details need to be included for different kinds of study. Following these guidelines can bring the following benefits:

- Increased confidence that you are reporting the right details of your studies. You can refer to the reporting guidelines you are using in your report to head off peer review requests to add unreasonable extra detail.
- It can make it quicker and easier to write up your reports, drawing on the guidelines to help structure them.
- Referring to guidelines can also be useful when you are planning your study. By making sure that you are considering all the details you know you will need to report, you can avoid missing doing or recording anything important during research design and collection.
- Referring to reporting guidelines demonstrates your commitment to best practice in TReQ research.

Why not?

Reporting guidelines are often developed for a very specific purpose, such as medical randomised control trials. Even if you identify the most suitable type of guideline for your own research, they may still ask you to report details that just aren’t relevant to your case. If so, you should feel free to exclude them – but it is better to consider and exclude them than not consider them at all, and risk missing out reporting on important points.

Following guidelines can be a bit time-consuming, for example if you have to create a diagram of participant flow through a study. However, there can also be time savings in making it easier to decide what to include/exclude.

For what type of research?

Reporting guidelines are available for many different research approaches (see How section below). Some are more detailed and prescriptive than others. Generally, the guidelines for reporting qualitative research are somewhat more general and flexible than those for, for example, randomised control trials. When in doubt, find the guidelines that seem most relevant to your work and use them as a starting point.

When in the research process?

Reporting guidelines are, as the name suggests, most commonly referred to at reporting stage (e.g. when you are writing a paper). However, it can be helpful to familiarise yourself with them earlier on in your project planning. This is because they provide a clear list of the important considerations in research design, and therefore provide a checklist that can give you confidence that you are covering all the bases.

How?

The table below provides links to reporting checklists for some common types of study. Along with the checklist, there is often a lot of information about how the checklist was arrived at and how it should be used.

Study type	**Reporting guidelines**	**Notes**
Randomized trial	CONSORT	Requires flowchart of phases of trial and includes 25-item checklist.
Systematic review	PRISMA	Flowchart and 27-item checklist.
Predictive model	TRIPOD	22-item checklist.
Qualitative study (interviews and focus groups)	COREQ	32-item checklist.

Further options and a guide to which checklist to use are available at Equator Network.

You can either use the checklist to explicitly structure your report or, at the very least, go through your report and make sure that all the (relevant) details are included somewhere. Using specific headings can make it easier for readers to locate specific details in your report. It is usual to state (with a reference) in your report which guidelines you are following. This makes it clear that you are committed to best practice in reporting research, and helps track how reporting guidelines are being used.

Further reading

Equator Network: "Enhancing the quality and transparency of health research", contains rationale for, and a library of, reporting guidelines

Amsterdam Public Health Quality Handbook: A further at-a-glance list of guidelines with links.

Consumer demand for time of use electricity tariffs: A systematized review of the empirical evidence: Example of a paper in energy which uses reporting guidelines (in this case PRISMA for systematic reviews).

Systematic evidence review

- A systematic approach to reviewing evidence involves following a pre-established procedure, documenting the review as it progresses, and reporting full details of how it was conducted.
- It you and others confidence that you haven't missed important evidence.
- It makes it easier to update or reproduce the review, and for you and others to highlight specific strengths and weaknesses.

Why?

The starting point of much research is a review of the existing literature. We want to discover what is already known with regard to the questions we are interested in, and where the knowledge gaps are. However, depending on how we go about identifying the existing literature, we might miss important existing evidence. This can result in wasted effort through duplicated research (e.g. if you perform a small study to test a hypothesis that larger studies have already disproved), or just a missed opportunity to draw on the most relevant research to inform your work and make it as insightful as it could be.

Employing a systematic approach to evidence review can give you confidence that you are not missing any key existing work, and increase the chances of you finding material that will help improve your own study. By 'systematic approach', we mean following a pre-established procedure, formally documenting the review as it progresses, and reporting full details of how it was conducted. This has the additional benefit of transparency – others can assess the rigour of your review based on the details provided. Depending on the kind of review, it can also allow easier updating – as any updating you or others perform does not need to go back and re-review the original material (so long as you have confidence that the original review was rigorously performed).

Perhaps the key benefit of systematic approaches to evidence synthesis is that they give us a much better representation of what is known about a given topic than any individual study ever could. They allow us to assess the 'weight of evidence' across many study types with different strengths and weaknesses and in different contexts, while minimizing risk of bias in terms of what evidence is included. For this reason they are particularly valued by policymakers who want to be as sure as possible that policy measures they introduce are informed by the best and most relevant evidence.

The main benefits to you from using systematic approaches to reviewing the literature are:

- Confidence that you are not missing important previous work.

- Greater knowledge of the topic in question and hence the ability to design a better study
- The ability to demonstrate the rigour of your review.
- In the context of a PhD thesis, it makes it easier to turn your literature review into a journal paper since you can more easily defend it against peer reviewer critique.
- The ability to seek early input on the plan for your review, based on your review protocol (see How section below).
- You can easily update the review as the project progresses without having to waste time going back over old ground.
- It can help familiarize you with different terms employed in your field as studied by other disciplines.

Why not?

The process of both devising and conducting a systematic-type review can seem daunting and time-consuming. However, compared to a more informal approach to reviewing, it can actually save time in the long run since it reduces the cycle of searching for literature, identifying useful documents, then forgetting which searches you have run and running them all over again. The process also gives much more comfort that key evidence is not being missed, which can be a big stress eliminator!

For what type of research?

It is possible to systematically review all kinds of research evidence, and systematic review approaches can be applied in any area. The specific approach to the conduct, analysis and reporting of the review will differ depending on the kind of evidence being reviewed. For more details see the How section.

When in the research process?

Reviews are most usefully completed early on in the research process, since then they can inform subsequent development of research questions and research design. However, the review should be periodically returned to throughout the work. This is because new evidence could have emerged since the original review was conducted, but also because you may have become aware of new strands of work which were not captured in your original review, and should now be included. And of course, systematic reviews often constitute standalone research projects in their own right.

How?

Systematic reviewing is a research method in its own right and substantial books have been written on the topic (see Resources section below), so we can only scratch the surface in this document. However, we are including some of the most

important details so you know basically what is involved. Firstly, there are a range of different kinds of systematic review, including:

- Full systematic reviews, which 'seek to collate all evidence that fits pre-specified eligibility criteria in order to address a specific research question' (Higgins & Green 2008). Such reviews are exhaustive and, therefore, resource intensive.

- Meta-analyses, when you statistically analyse the findings of previous work in an attempt to (for example) show the effect size of an intervention across multiple studies.

- Evidence maps, which describe (usually at a very high level) the nature and extent of research on a particular topic, and help identify either gaps or areas for subsequent more detailed review.

- Rapid evidence assessments, which take a similar approach to full systematic reviews by which alter the review criteria to reflect resource (e.g. time and money) constraints. For example, they might be restricted to publications after a certain date, or only in certain countries.

- Realist synthesis, which focus more on how particular outcomes have come about and how this depends on the specific context in which interventions were carried out.

Whichever review you choose you want to first check around and see if anyone has done a review on it before. Some subjects have registers of systematic reviews which are in progress (see e.g. the Collaboration for Environmental Evidence, or PROSPERO for health studies). You will also need to do some pilot searches to see if there is enough (or too much) material for you to cover. This will help you develop a specific review question (or questions).

It is then standard practice to write and, ideally, publicly share, a review protocol. This sets out the way in which you will conduct you review and, similar to other forms of preregistration (see preregistration section above), can give readers more confidence in your findings and help you get early input from advisors. The stages of the review that you will need to include are:

- How you will identify material to include. This will likely involve specifying keywords to search bibliographic databases, broader search approaches (such as of company websites), backward/forward citation tracking, and/or calling for evidence from other researchers. Be aware that in applied areas of research, the most valuable evidence is as likely to be found in grey literature (e.g. government/NGO reports) as in academic journals.

- Your inclusion/exclusion criteria. That is, of all the material you identify, how you will choose what to review. This might involve specifying things like only including empirical evidence, or only material published after a certain date, and (of course) specific subject focus relevant to your review.

- What you will extract from the material you review. This can be anything you need to answer your review question(s).

- Quality assessment. In your final review you want to give more weight to better studies which are more relevant to your question of interest. A range of pre-existing tools are available to help you make these assessments.

- Synthesis and analysis. How will you actually bring together the evidence you have identified and being to draw conclusions from it? This may involve doing statistical analyses, thematic re-analysis of qualitative data, or more narrative approaches.

- Reporting. Guidelines such as PRISMA are available to let you know which are the most important details of your review to report.

Much more detail on conducting systematic-type reviews is available in the resources below.

Further reading

The UKERC Technology and Policy Assessment programme has conducted many reviews with systematic approaches in the subject of energy, along with a guide to developing a rapid evidence assessment (available at the link).

Civil Service guidance (now archived) on rapid evidence assessment.

Example of a rapid realist review protocol.

Acknowledgements

Both authors contributed equally. This work has grown out of an original paper we wrote for the ECEEE summer study conference 2017. We would like to acknowledge the contributions of all the co-authors -- in particular Moira Nicolson (who was centrally involved in writing the paper) and Harry Kennard (who presented the results at other conferences). We would like to thank Nicole Watson and Harry Kennard for their very helpful feedback on a draft of this document.

Michael Fell and Gesche Huebner are supported by the UKRI Centre for Research into Energy Demand Solutions (grant number EP/R035288/1).