

# Research rigour, transparency, and reproducibility in energy research

The reproducibility crisis refers to the fact that many established research findings cannot be reproduced; [over 70% of scientists asked by Nature](#) said they had been unable to reproduce other groups' work (see note 1).

End-use energy demand research is very multidisciplinary; we find engineers, architects, mathematicians, economists, sociologists, anthropologists, statisticians, and others researching this field using a multitude of methods, such as interviews, focus groups, surveys, field and lab experiments, case studies, monitoring and modelling.

End-use energy demand is a crucial research field given the imminent need of reducing our carbon emissions. So, one might argue that we should deliver the most robust, transparent, and reproducible research – yet, we are not as far advanced as other, more homogenous disciplines when it comes to reproducibility. Why? (see note 2)

The multitude of methods that we use likely plays a role – for many of those, 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 (see note 3).

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 transferable 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.

But in all areas of our research, we should improve transparency and rigour, being clear about what we do and being better in how we do it. This could mean

- Pre-specifying how we will collect and analyse qualitative data.
- Publishing our coding frame and coding annotations for qualitative data.
- Making clear how we arrive at modelling assumptions and what data sources we use.

- Publishing as much data as possible and being bolder in negotiations with external bodies on making data available.
- Focusing on systematic reviews and using reporting guidelines
- Learning from other disciplines about their best practices for a specific method.

And of course, for a substantial part of our research, the more quantitative, empirical, experimental type, we can use common tools such as preregistration, i.e. specifying our analyses before getting our hands on the data and uploading such plan to an online repository (see note 4).

More details on tools for aiding rigour, transparency, and reproducibility will follow in a separate blog post (see also [a related conference paper from eceee](#)) and on an online platform that Mike Fell and I are developing (hopefully with support from the energy community!). If you are interested in contributing, do get in touch with [Mike](#), @mikefsway or [myself](#), @GescheHuebner).

## Notes

1. I differentiate (following the reference in footnote 1) between reproducibility and replicability as follows: 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.
2. There are lots of cultural and institutional issues such as the pressure to publish and the emphasis on novelty for funding; these apply to many disciplines and I won't discuss them here.
3. The [Smart Energy Research Lab](#) (SERL) currently develops a 'safe' research portal to give UK based researchers access to half-hourly smart meter data with some contextual information.
4. This practice contributes to overcoming (1) the 'file drawer' problem, whereby results that do not exceed conventional thresholds for

statistical significance are less likely to be submitted or accepted for peer-reviewed publication (and hence remain invisible); (2) 'fishing', whereby researchers select analysis specifications that support their prior beliefs, desired conclusions or simply give a significant result.