

I'm not surprised

It is easy to conflate what is known based on the scientific literature and what feels known because it is intuitive. However, empirical validation and precision are particularly critical for policy-relevant behavioural research, regardless of whether the results are surprising.

Predictions about how two chemical compounds will react are based on knowledge of their chemical properties and scientific theories that underpin our understanding of chemical reactivity; a surprising result would therefore be one that is unanticipated based on the accepted scientific literature. Drawing predictions based on previous observation and/or theory is the basis for the scientific method across disciplines, but inference from educated intuition is equally pervasive. This poses unique challenges for studies of human behaviour, where subjective experience provides an additional source of intuition. A microbiologist, for example, can have intuitions of how bacteria should behave under certain conditions based on accumulated evidence, but these can never be based on their subjective experience of being a bacterium. In contrast, human behavioural results can be seen as surprising when they violate what people believe about their own behaviour, despite decades of research showing that people do not have good insight into why they behave the way that they do¹. Because surprising and counterintuitive findings subjectively feel more interesting, they tend to be rewarded and garner the most attention.

Incentivizing publication of surprising findings and related publication biases that discourage null results and replications creates problems that are not new^{2,3} (and not limited to the social sciences^{4,5}). However, these problems are particularly acute when considered in relation to policy-relevant research, where empirical validation and quantification are more important than discovering something unexpected. For instance, in this issue, Moira Nicolson and colleagues (article no. 17073) demonstrate that unsolicited mass e-mails are effective for promoting time-of-use energy tariffs, particularly when they target a specific behaviour (in this case, electric vehicle charging) and are optimally deployed within the first three months of electric vehicle purchase. This provides evidence that an easy-to-implement, low-cost intervention could, in the authors' estimation, effectively engage an additional one million people with time-of-use tariffs once electric vehicles

reach 60% market penetration. This is information a policymaker can use. However, in the absence of this behavioural evidence, policy decisions could be subject to the whims of a policymaker's assumptions about how they think they would behave in similar circumstances or how they have behaved in the past. A policymaker could dismiss a proposal for such an e-mail scheme because they would never open such e-mails, or endorse the proposal because they routinely open promotional e-mails. Critically, the results of research like that of Nicolson and colleagues are important regardless: either the data are surprising and violate assumptions, or they provide empirical confirmation that intuitions are valid. Either outcome is significant when evaluating a potential policy or programme that will entail real costs to deploy and that is being implemented because there is a problem that needs to be solved.

Nicolson and colleagues tested their e-mail intervention in the UK. Beyond practical circumstances that may limit implementation in other countries, should we expect equivalent success for a similar intervention in the US or in China? National context is not generally considered in the biological or physical sciences: the basic principles of physics do not vary around the world. However, behaviour and the factors that influence it are highly dependent on complex socio-cultural factors. Consequently, energy use behaviour and the effect of conservation policies differ between countries, and not simply because of differences in exogenous factors like climate (which influences energy needs for heating or cooling) or development (which limits energy options)^{6,7}. Theoretical framework determines whether it is interesting to compare energy use patterns or intervention effectiveness between two countries (and whether finding similar effects is surprising), but the empirical result is important for policy decisions in the countries studied regardless.

Luckily, the opportunities for conducting policy-relevant behavioural research are poised to expand in unprecedented ways. In a Comment in this issue (article no. 17085), Verena Tiefenbeck highlights how behavioural and social sciences could harness

information technology advances to better understand energy consumption behaviour. The growing ubiquity of smartphones and smart meters, and the falling costs of sensors, communication infrastructure, processors, and data storage, foreshadow a gold-mine of data that researchers could use to track real-world real-time energy-use behaviour in large, diverse, representative populations over long periods of time. Collecting moment-to-moment data on daily behaviour will likely generate a lot of unsurprising intuitive findings (for example, larger households consume more energy). However, the promise of big data is not necessarily new revelations about mechanisms, but precision of measurement. In basic research, it is sufficient to show that a manipulation produces an effect, which can inform theories about underlying mechanisms; in contrast, to justify policy implementation, it is essential to know how big that effect is, with precision, to ensure that the benefits justify the cost.

In a recent episode of the Forecast Podcast, Solomon Hsiang (University of California, Berkeley) told *Nature's* climate science editor Michael White that "Doing economics is like trying to derive $PV = nRT$ [the ideal gas law] when you are one of the molecules inside the bottle" (<http://go.nature.com/2rKjLhc>). This sentiment applies to the social and behavioural sciences at large, where people are the subjects they study. There is no easy solution to this problem and it is easy to dismiss a result as uninteresting because it is 'obvious'. But authors, reviewers, and editors can be more aware of the biases they bring to designing studies, interpreting data, and evaluating manuscripts; we can step back and question why we do not feel surprised by a result, and then decide whether being surprised is important. □

References

1. Nisbett, R. E. & Wilson, T. D. *Psych. Rev.* **84**, 231–259 (1977).
2. Nosek, B. A., Spies, J. R. & Motyl, M. *Perspect. Psychol. Sci.* **7**, 615–631 (2012).
3. Smaldino, P. E. & McElreath, R. R. *Soc. Open Sci.* **3**, 160384 (2016).
4. Ioannidis, J. P. A. *PLoS Med.* **2**, e124 (2005).
5. Ioannidis, J. P. A. & Khoury M. J. *Science* **334**, 1230–1232 (2011).
6. Wilhite, H., Nakagami, H., Masuda, T., Yamaga, Y. & Haneda, H. *Energy Policy* **24**, 795–803 (1996).
7. Ma, G., Lin, J., Li, N. & Zhou, J. *Energy Build.* **134**, 329–338 (2017).