

World view



By Paul Smaldino

Better methods can't make up for mediocre theory

With better questions, many reproducibility problems will fall away, says Paul Smaldino.

Much digital ink has been spilt describing ways to improve replicability in science. Preregistration. Open data. Open code. These are all necessary, but insufficient. The thing is, we don't just want science to be reproducible. We want it to help us to make better sense of the world.

For that, we must create better hypotheses – and those require better models and better measurements.

A theoretical model of mine (P. E. Smaldino and R. McElreath *Soc. Open Sci.* **3**, 160384; 2016) made headlines when it showed that bad science – or rather, less rigorous science that could produce more papers in less time – could crowd out the more robust sort. This suggested that generating better hypotheses is at least as important as reducing methodological errors for minimizing false discoveries.

Who cares if you can replicate an experiment that found that people think the room is hotter after reading a story about nice people? Will this help us to develop better theories? You can craft a fun story about that result, but can you devise the next great scientific question?

To generate good hypotheses, we need good theory. In a landmark study attempting to replicate 100 psychology papers, cognitive-psychology studies were replicated about twice as often as those from social psychology (Open Science Collaboration. *Science* **349**, aac4716; 2015). I think that's because cognitive psychology has better theories.

Good theory has at least two requirements. First, it can be used to build mathematical or computational models that derive clear, testable consequences from our assumptions. Every mature scientific discipline has these. Physicists use models of force and momentum to predict the motion of materials. Epidemiologists use models of contagions to understand the spread of disease. Neuroscientists use models of neural-spike trains to understand information flow in the brain. Social scientists use game models to understand the emergence of social norms.

Second, good theory must make sense, or at least acknowledge its contradictions. Consider the 'pre-cognition' studies of US social psychologist Daryl Bem, which were completed with remarkable transparency (D. J. Bem *J. Pers. Soc. Psychol.* **100**, 407–425; 2011). (The general consensus is that these studies did not establish the presence of extrasensory perception in college students, but the prevalence of overly flexible statistics; Bem defends the statistics as sound.) The work flouted well-supported ideas about physics and causality. It was akin to when physicists at CERN, Europe's particle-physics laboratory near Geneva, Switzerland, 'discovered' faster-than-light neutrinos, violating the special theory of relativity. Because the researchers required

We don't just want science to be reproducible. We want it to help us to make better sense of the world."

Paul Smaldino is an assistant professor of cognitive and information sciences at the University of California, Merced. psmaldino@ucmerced.edu

their results to be consistent with a broad theoretical framework, they probed deeper and discovered that their finding stemmed from a loose fibre-optic cable. To be clear, it's not the case that surprising claims are always wrong – but such claims must undergo extensive scrutiny.

If useful models produce better science, then what drives better models? Improved measurements. Consider the work of Tycho Brahe – a great astronomer of the sixteenth century, who nonetheless thought that the Sun orbited Earth. Yet his painstaking measurements of the positions of the planets allowed Johannes Kepler to determine that their orbits are elliptical. From this, Isaac Newton could formalize his theory of universal gravitation, which allowed modern researchers to ask countless questions about planetary motion, cosmology, ballistics, engineering and more.

If we can't reliably measure something, it's hard to build a theory about it. Quantities such as position, mass and time are relatively easy to measure, at least at some scales. Cognitive scientists can readily measure skin conductance, reaction times and word counts; this allows regularities and variation to be observed, and thus the construction of testable models. Other fields, including those I work in, have struggled with measurements. Psychologists attempt to measure emotions, identities and beliefs. Social scientists attempt to measure inequality, polarization and disinformation. Biomedical scientists attempt to measure treatment outcomes in small, heterogeneous populations.

I think that many sciences struggling with replication are those with the most pressing challenges in taking clear measurements. The trick lies not in merely finding a measurement that can be made precisely or described transparently, although these factors are important. Instead, scientists must find properties that can be reliably measured, inform theory and lend themselves to quantification in formal models.

Ideally, strong theories, formal models and measurements will interact in a virtuous cycle. Models allow us to study assumptions about the world and discover their consequences. The results can show what measurements are needed to test the assumptions, and those measurements can provide empirical patterns that invite explanations, which models can provide. And on and on.

We absolutely need better methods for hypothesis testing, and these are already being incorporated into how scientists are trained and how science is done.

So now it is time to focus on better practices for hypothesis generation. We need training programmes in model building and critique, plus consortia-building and funding programmes to invent and test measurements that make models tractable.

Better methods will help us get the right answers; models and measurements will ensure we ask the right questions.