# Reality Checks for Academia

*Bradford Tuckfield*

There's an old story about a shipbuilder who takes on a young apprentice. The apprentice learns some of the basics of the shipbuilding trade, but is careless and takes shortcuts that threaten the seaworthiness of the first ship he builds. After the ship is complete, but before it has been tested on the water, the apprentice asks the master shipbuilder for forgiveness for his mistakes. The master says that of course he can forgive him but the sea does not forgive. The reality of physics treats all ships equally regardless of who made them and whether their mistakes were forgivable; they will simply float if they are seaworthy and sink if they are not.

Putting a finished ship on the water provides a reality check for any shipbuilder. Indeed, the whole shipbuilding trade is disciplined by this final test. Charlatans who propose Swiss-cheese ships filled with holes can easily be repudiated, and the body of theoretical shipbuilding knowledge is likely to be entirely correct thanks to the harsh discipline of reality.

The scientific research done at institutions of higher education ostensibly has a connection to reality comparable to the connection between the theory and the practice of shipbuilding. Medical researchers who strive to create cures for life-threatening diseases must face the reality check of whether their proposed cures actually work. Like checking whether a boat floats or sinks when first placed on the water, doctors can check whether and how patients respond when given a proposed treatment. Academic engineers must pursue ideas that lead to cars that drive, planes that fly, and bridges that stay up. Computer scientists must publish algorithms that perform their stated functions. In each case, reality checks discipline scientific fields to remove charlatans, frauds,

*Bradford Tuckfield is a data scientist in Phoenix.*

and honest mistakes — at least in theory.

In practice, recent failures and crises in the world of academic research have exposed deep and systematic problems in how scientific research is conducted and evaluated, and how academia operates in general. These problems should not only be the concern of academics. Government funding of academic-research institutions is large enough that any taxpayer should want academia to function properly so that his tax dollars are not wasted. More seriously still, academic research and higher education can have huge spillover benefits that make the proper functioning of academia a matter of national importance. The lack of reality checks in academia has led to a series of costly failures, and the future of academic science rests on reasserting the quality control that reality provides.

## The Replication Crisis

The most notorious problem in academia today is the so-called "replication crisis." This embarrassing predicament has been widely covered in major media outlets such as the *Atlantic*, *Slate*, and *New York Magazine*, as well as on popular blogs run by prominent scientists Uri Simonsohn and Andrew Gelman.

Consistent replication of scientific findings is crucial for a properly functioning scientific enterprise. If a research finding is true, like the finding that ships with holes always sink, it should be replicable by researchers who are independent of those who published the original finding. If a research finding is not replicable, then it is not "true" in the scientific sense, since science is concerned not with telling stories, but with discerning the general laws that govern the universe. If an alleged phenomenon does not replicate, we must conclude that the universe functions in a way that does not cause it to replicate, or in other words that the finding is somehow rooted in error. As Wharton professor Joseph Simmons recently wrote in a blog post, "History is about what happened. Science is about what happens *next*." Every serious scientist agrees at least in principle that replication is important.

If even one scientific finding is accepted as true but is not replicable, it can cause serious damage in a few ways. First, since science is a cumulative effort, new research builds upon old research. If the old research is false, the new research will be useless at best and a deceptive waste of time at worst. Second, a non-replicable research finding can





lead to career rewards for researchers who are charlatans or frauds, or who make honest but serious mistakes. One might imagine that when charlatans or incompetents are given responsibility over prestigious departments, journals, or institutions, their charlatanry or incompetence is amplified and reproduced.

Finally, a non-replicable finding that is promoted in the popular press as true can mislead and harm the millions of laypeople who hear about it. Discredited research cited by anti-vaccination activists is a particularly egregious example of this. In each of these ways, even one non-replicable research paper can cause serious and snowballing harm.

Unfortunately, the current scientific literature contains not just one, but a huge number of findings that have not been generally replicable — that is, when independent researchers follow the methods outlined in published papers, they do not obtain the results that the original paper claims. The term "crisis" is no exaggeration: The inertia of the enormous world of academia and its traditions, the unwillingness of some to act or make helpful changes, the nonzero rate of malicious fraud, and the enormous potential negative consequences, together make this a truly daunting problem that will have long-term consequences. The complexity of the issue has also made it difficult to isolate its causes, and therefore more difficult to solve.

Given the universally acknowledged importance of replicability, it is a wonder that the world's top scientific journals have published dozens or even hundreds of non-replicable research findings, and that their non-replicability has gone undetected for so long. The practice of replication, enshrined in good science for centuries, should act as a reality check just like putting a new ship on the water. Why has this reality check failed, and how can we improve it, restore it, or replace it with something that will properly discipline scientific research and ensure its robust functioning?

## The Center of the Crisis

One reason that reality checks can fail in scientific research is that the improvements proposed in new research are often *small* or *probabilistic*. For example, a scientist might be trying to improve a cancer drug that has been shown to improve by 52% the chance of long-term survival, and his breakthrough leads to a 60% chance of long-term survival. That would be considered a significant advance, but it would not be

immediately obvious whether his drug truly represents an improvement in the way a solid ship hull represents an improvement on a Swiss-cheese ship hull. Rather than relying on a simple, obvious test like whether a ship floats, those who evaluate these drugs will have to rely on careful measurement, experimental trials, and statistical analysis to detect these small, probabilistic differences.

Experimental trials and statistical analyses are wonderful things, and the world could benefit from more of both. However, they are also difficult to properly conduct and draw correct inferences from, especially when the studied effects are small. The necessity of complicated trials and sophisticated statistical analyses for the evaluation of a researcher's claims makes it harder for research programs to benefit from reality checks. Some honest mistake or malicious fraud in the performance of a clinical trial or statistical analysis could make true breakthroughs seem worthless, or make worthless ideas seem like true breakthroughs. Eventually, the truth should win out, but in the absence of strict, simple reality checks, it can take many years for people to realize that research programs are founded on mistakes. In the meantime, researchers can get tenure based on these mistakes, and, more seriously, drugs can be mistakenly approved or rejected and damage people's health.

Other scientific fields have comparable potential problems. An academic engineer might design a bridge that has a 0.03% chance of collapsing over a given period of time instead of a 0.05% chance. An academic geologist might devise a method to predict earthquakes with 65% instead of 60% accuracy. In each case, mistakes in experimental trials or statistical analyses could lead to charlatans getting tenure and other career rewards, and to research literature being corrupted by false ideas and failing to incorporate correct ones.

Some fields, though, simply have looser connections to physical reality, making reality checks less effective. *Slate* recently published an article by well-known statistician Andrew Gelman stating that the replication crisis seemed to be worse in psychology, and especially in social psychology — a field that benefits relatively little from rigorous reality checks.

To say that psychology is more loosely connected to reality than other fields is not meant as a mean-spirited critique. Most social-psychology practitioners would admit that the field is heavily focused on small, probabilistic effects. For example, one well-known social-psychology





study purports to demonstrate "social priming" by exposing subjects to words related to the elderly, then surreptitiously measuring how fast the subjects walk. Researchers tested the theory that merely reading a word associated with elderly people will trigger mental associations with the elderly, which will manifest themselves in slightly slower walking speeds (since one stereotype about the elderly is that they move slowly). The claimed effect was quite small (a difference of about one second of walking time over a fixed distance) and probabilistic (it didn't happen every time, to every subject, in a predictable way, but only increased the chance of walking more slowly).

Effects that are so small and so dependent on chance and averages are not easy to confirm or falsify; they can be studied only through meticulous trials and careful statistics. When those trials or statistical analyses fail (in the sense of failing to reject a falsehood or confirm a truth), scientists may begin to believe things that are not true. This can corrupt entire scientific disciplines.

There is another important reason why social psychology's connection to reality is relatively weak. Unlike many other scientific fields, social psychology rarely leads to the creation or improvement of technology in the way that theoretical shipbuilding leads to building ships. In social psychology, breakthroughs tend to happen in labs and stay there, and so these ideas do not face the rigorous and unforgiving test of whether technology based on them functions.

## Technical difficulties

Replication is one important reality check for academic research. Another is the development, testing, and deployment of technology. Corporations that develop technology based on research findings have a profit motive, so they seek replicable, reliable research with meaningful effects, since that is the research most likely to lead to profit. The free market has its own disciplinary mechanisms that tend to punish technology based on false findings and reward technology based on true ones. Academic research that is closely connected to corporate implementations of technology naturally benefits from these market-based reality checks.

Perhaps the most obvious example of research leading to technology is found in medicine: Basic research in chemistry is followed by preclinical medical or biological research, then the academic development of

drugs, then research by for-profit pharmaceutical companies, and finally the deployment of drugs in the mass market. Another example would be engineering research in academia, where researchers seek theoretical breakthroughs that can be translated into engineered technology by corporations (making cars, planes, ships, and many other things).

Most scientific fields have a "supply chain" through which abstract and theoretical ideas are translated to more and more practical applications and finally to technology that is used in non-academic contexts in the "real world." Academics sometimes justify their theoretical work by referring to its eventual benefits to downstream technology, but in fact, by forcing academic ideas to face reality checks and be disciplined by reality, technology also has an upstream benefit for those who are developing the abstract ideas.

Technology based on research need not be physical. Research in economics and finance frequently leads to technologies like investment vehicles and strategies and insurance-contract structures. One example of a failed economic technology was the hedge fund Long-Term Capital Management (LTCM). This fund recruited prominent economists Myron Scholes and Robert Merton around the time of its founding to provide advice on the macroeconomy and investment strategies. This fund was a "technology" in the sense that it was an application of theoretical ideas from academic research, applied in actual functioning markets to create value. In 1997, around the peak of the fortunes of LTCM, Scholes and Merton shared the Nobel Prize in Economics for the types of ideas they were providing to LTCM.

Then in 1998, LTCM experienced a disastrous collapse and lost billions of dollars — nearly its entire portfolio. In retrospect, investors would have done much better by investing in U.S. Treasuries or keeping their money in a safe than by investing with LTCM over its history. As it turned out, the ideas that LTCM based its strategy on were great on paper, but were unable to withstand the chaos of the real world.

Funds fail every day, and though LTCM's collapse was very large, it was not an unheard-of loss. Worse than the failure of the fund, however, has been the failure of academic researchers to incorporate this error into their worldview. LTCM was a reality check: Two Nobel Prize winners put their ideas to the test in a real fund with real dollars in a real market. Their brilliant theories failed catastrophically in practice. At the very least, we should repudiate the ideas that LTCM tried to implement.





Unfortunately, the reaction in academia to the failure of LTCM was not only small, it was virtually nonexistent. There should have been a great increase in skepticism toward these scholars and their ideas, but there has not been. They maintain prestigious academic and honorary corporate positions, and their papers, textbooks, and mathematical models are still widely read, taught, cited, and taken seriously around the world. The failure of the fund is nothing compared to this failure of an entire field to learn from an obvious reality check.

LTCM was like a ship that looks beautiful as it leaves the harbor, wins the shipbuilder many plaudits and awards, and then sinks. One might assume in such a case that the shipbuilder's reputation would suffer. Yet Merton and Scholes are still revered in academic economics and finance. This unfortunately vindicates some of the worst accusations against academia. Though we may be impressed by elegant math and brilliant minds and are willing to forgive mistakes, the financial markets (like the sea) do not forgive, and we should take their verdicts seriously.

## BEYOND SOCIAL SCIENCE

The cases above concern social psychology and economics, both social sciences. Naysayers sometimes dismiss these "soft sciences" as inherently less rigorous than the "hard sciences" such as physics, chemistry, biology, geology, and the like. If replication crises and technology failures occurred only in the social sciences, that could provide evidence that they are indeed "soft," and inferior to physical sciences. However, these "hard" science fields have had serious problems of their own.

Cancer research is one high-profile, enormously expensive undertaking that is beginning to confront its own deep problems. While these could fill a book, its greatest challenge is essentially identical to social psychology's: a failure to replicate small, probabilistic findings. While some cancer researchers must contend with the reality checks of technology testing and adoption, a great deal of "preclinical" cancer research takes place far away from these real-world constraints, and is therefore less disciplined by reality checks — meaning non-replicable findings can more easily slip through unnoticed. In addition, the cost of replicating experiments — or really, conducting any experiments — can be prohibitive. It may therefore be difficult or impossible to raise enough money to double-check scientific studies, allowing flawed research to stand

unchallenged.

None of the physical sciences is immune to these problems. Research in chemistry, physics, biology, and other areas of medicine has yielded findings that cannot be replicated and that have caused insidious damage. And in fields where experimental trials are usually not feasible, such as epidemiology and climate science, research can suffer from the lack of these useful reality checks. While studies in those fields often have reasonable descriptive accuracy, they frequently fail to properly identify causal models of the phenomena being examined.

The humanities in general do not rely on empirical checks or experimental trials of proposed ideas. Nor do new ideas in the humanities typically lead to technology and its attendant reality checks. Rather, human judgment, wisdom, and intuition are the yardsticks by which truth is measured. To the great credit of the humanities, they have produced many excellent ideas and insights without the reality checks that accompany the hard sciences.

However, that is not to say that humanistic scholarship has never gone astray or lost touch with reality. In a recent *City Journal* article, philosopher Roger Scruton pointed out a few cases in which humanities scholarship seems shockingly harebrained. Scruton's criticisms were mostly of Slavoj Žižek and Jacques Lacan, two lionized sages with undoubtedly brilliant minds, whose writings are nevertheless impenetrable (Scruton uses the apt term "logorrhea"). Here is one of the Žižek passages that Scruton cites:

> Is not the paradoxical topology of the movement of capital, the fundamental blockage which resolves and reproduces itself through frenetic activity, *excessive* power as the very form of appearance of a fundamental *impotence* — this immediate passage, this coincidence of limit and excess, of lack and surplus — precisely that of the Lacanian *objet petit a*, of the leftover which embodies the fundamental, constitutive lack?

It is certainly possible that Žižek and his ilk possess knowledge that enables them to communicate in ways that are incomprehensible to the average college-educated citizen. Or, it could be that the emperor really has no clothes and Žižek's oeuvre is filled with nonsense. Scruton, who is highly intelligent and educated, and who has engaged seriously with





Žižek's writing, favors the latter explanation. The prevalence of this type of claptrap in the humanities indicates a drift away from meaningful engagement with reality and toward a Žižekian fakery with little or no connection to the truth.

Fields like "critical theory," gender studies, queer theory, and other recent inventions offer up some truths and some nonsense in the style of Žižek and Lacan. It is beyond the scope of this essay to delineate where the truth ends and the nonsense begins. Unfortunately, to the extent that nonsense exists in these fields, it is harder to root out than it would be in the empirical sciences, since truth claims must depend on appeals to intuition and judgment rather than on commonly accepted and objectively verifiable tests.

## Toward Some Reality Checks

Scientists and various friendly observers of the academy have proposed many solutions for the replication crisis and academia's drift from reality, many of which boil down to conducting better experimental trials and statistical analyses. These are reasonable ideas, but they are not enough. Statistics and experimental trials will always be hard, even for experienced scientists, and both honest mistakes and fraud will continue to sneak past the gatekeepers (journal editors, peer reviewers, and so on) and into the scientific literature.

While there is no silver bullet for these complex problems, a salutary first step would be to push academic research to have a close connection to reality, which would push science to face strong and robust reality checks. As noted above, even scientific research can avoid the discipline of reality by studying small or probabilistic effects, or by not connecting to some downstream technology. Research agendas should therefore be guided toward technologies, and researchers should focus on large, meaningful effects. Despite these general principles, though, reality checks will look different in each research field, as they each have different strengths, weaknesses, and challenges.

Social psychology, for example, would benefit from focusing on large, meaningful effects, instead of small, probabilistic effects that are plagued by statistical traps. One such large research focus could be human happiness, a conundrum that people earnestly seek and struggle to reliably untangle. If the psychology profession believes that it can successfully plumb the depths of the human mind, then it should be able

to more successfully take on the question of what makes people happy. This type of research would benefit from the reality check of widespread public scrutiny, as people will always be interested in happiness. And since the difference between happiness and sorrow is so vast, there are presumably large effects to be studied. This type of research would be both more meaningful and have a greater chance of success than probabilistic tests of small changes in walking speeds based on stereotypes.

Social psychology can also better connect itself to technology, broadly understood. A great deal of social-psychology research is conducted by business-school professors, including professors in marketing departments. Marketing and advertising are technologies that could provide tests of the ideas of social psychology. Today, the implementation of academic psychological ideas by corporate marketers happens haphazardly and independently of the scholarly publication process, if it happens at all. If social psychologists worked more closely with corporate marketers, they could benefit from the reality checks of the free market. We could then judge psychology papers not only by their reasoning and statistics, but by their impact and success in corporate-advertising campaigns. This would put at least part of social-psychology research more closely in touch with reality. Social psychologists whose work is not transferable to corporate marketing may be able to find other fields where they can influence technology. For example, social psychologists could conduct research that influences management practices, negotiation techniques, or pedagogy.

In economics, there ought to be a closer connection between academic research and technology. Economics professors, and especially finance professors, should subject their ideas to the discipline of reality by starting funds and making investments guided by their research. The funds and investments of these academics would constitute the "technology" by which their ideas could be simply and easily judged. University departments and the public could then study the performance of these technologies, rewarding those whose ideas lead to success and punishing (at least the careers of) those whose ideas lead to failure.

Of course, some academic economists and finance researchers study phenomena that are not readily convertible to technology. For example, a recent issue of the American Economic Review contained a research paper about the evolutionary or cultural origins of time preferences (e.g., the preference for one marshmallow now over three marshmallows in





five minutes or vice versa). It would likely be difficult for the authors of that research to concoct a fund or investment strategy based primarily on the ideas presented there. If such research has no ability to influence downstream technology, it should be considered less valuable than comparable research that does lead to technological improvements.

But more fundamentally, potential solutions to the problems of academia would involve overhauling or reforming academia's "infrastructure." For example, many have proposed the post-publication review of research papers. This would entail publishing every single academic paper (probably in an online repository where "publishing" doesn't face serious resource constraints), and then creating a moderated comments section where the author's peers could write public comments praising or criticizing the research. Rather than two or three anonymous and secret reviews, this could provide dozens of signed, open reviews. And in place of a handful of editors holding the keys to the world's top journals, anyone from anywhere could author and promote a serious breakthrough. Post-publication review would create a more democratic system for academia, and a reality check based on social consensus.

Other proposed changes to the university system include reforming the archaic master/apprentice system in doctoral programs, providing career incentives for replicating research, or abolishing tenure altogether. None of these are likely to be immediate panaceas, but something must be done; the academy's detachment from reality is crippling whole fields of research.

## REALITY AS A GOAL

In Plato's *Phaedrus*, Socrates describes the gods and other immortal souls as engaged in a constant chariot journey around the cosmos. The gods take a path that enables them to see the world of reality outside our universe, gazing on the true forms of Justice, Self-Control, Beauty, and other abstract realities. Other souls follow the gods, but are not always able to see these forms. Those who do are nourished by the sight — it helps their wings grow. Even one glimpse of reality enables a soul to continue its circuit and try to catch another look. Souls that complete a full circuit without viewing reality a single time lose control of their chariots and are incarnated on earth.

Socrates describes a hierarchy of human incarnations based on these

sightings of reality. Souls that have seen the most reality before coming to Earth are incarnated as philosophers, since they are closest to the divine. Those who have viewed the second-most reality are incarnated as righteous kings. The rest of the hierarchy goes as follows: politician, gymnast, prophet, poet, artisan, sophist, and tyrant. Through subsequent reincarnations, one attempts to climb the ladder to reclaim one's place in the circuit around the heavens.

Two lessons jump out at us from this peculiar metaphysics. First, Socrates puts the philosopher at the top of the hierarchy, while the sophist is only one step above a tyrant — the very lowest form of human. This gap may seem surprising, since both the sophist and the philosopher use rhetoric to persuade others of apparent truths. The difference is only in their experience of and commitment to truth. The philosopher is the soul who, according to Socrates, has had the greatest experience of true, divine reality. The sophist has had almost the least. The philosopher's profession, Socrates asserts, is to use rhetoric not for its own sake, but to convince others of the truth and the nature of the divine reality that he has experienced. The sophist uses rhetoric, but in a cynical or self-serving way to convince others of falsehoods or swindles.

Looking around academia, we can see some whom Socrates would describe as philosophers. They are dedicated to reality and make great sacrifices of time and energy to understand it. We can also see some whom Socrates would describe as sophists. They are skilled in the forms and conventions of academic research, but use them only as tools to gain money or prestige, rather than to approach and enjoy the truths of the cosmos. The effect of sophists on the landscape of academic research is to fill it with fakery: fake concern for truth, fake erudition and wisdom, and research that looks reasonable but doesn't replicate or otherwise correspond to reality. Somehow, we must ensure that the world of research is dominated by philosophers rather than sophists.

The second lesson of Socrates' story is that merely getting a glimpse of absolute reality is worth extraordinary effort. The souls he describes in the heavens have the view of reality as their primary goal, and one that may take thousands of years to accomplish. This devotion to truth is inspiring, and can lead one to rethink the petty motivations that often animate life. From this point of view, reality checks need not be a necessary evil. Rather, applying reality checks to research could become a sincere pleasure for scientists who love truth and are eager to find out





what reality can teach us.

Many motivations might move a scientific researcher in the course of a career. He may seek promotions and raises, tenure, or prestigious editorships and honorary positions. He may be motivated by the fame that occasionally comes to prominent scientists. He may be motivated by friendship, defending the theories of people he likes, or even spite, attacking the theories of people he hates. Sometimes a researcher is motivated by simple inertia, arguing for a discredited notion because he wrote a book or published a paper supporting it. It is unrealistic to think that scientists — who are, after all, human — could be immune to vain, petty, or misguided incentives.

What we can hope for is that the strongest, or at least a very powerful, motivation for academic researchers will be the pursuit and study of reality. No amount of statistics education, university reorganization, or even public-policy changes can solve the problems facing academia if its researchers are not earnestly and seriously focused on discovering the truth and presenting it honestly. Too often, this is simply not the case.

All is not lost, however. University hiring committees can think seriously about each hiring decision to ensure that sophists are kept out. Administrators can make funding decisions that reflect each department's connection or commitment to reality. Members of the public can become less credulous, and demand to see replications of alleged findings and technologies that successfully implement them, rather than simply believing every overhyped research finding. And researchers can examine their own motivations, becoming a little more concerned with the reality of the cosmos and a little less concerned with the reality of their next promotion, raise, or speaking fee.

A stronger connection between academia and reality would yield innumerable benefits: Millions of taxpayer dollars would go toward research that improves life, rather than dead ends, fakery, or non-replicable wastes. Followers of academic research would be more in touch with the truth, since academia would be better at weeding out falsehoods. And if American universities take the lead in solving the replication crisis and other problems in academia, they could attract skilled scholars and top students. We would also educate our own citizens better, and have a more complete understanding of reality. Or, to paraphrase Socrates, we would grow our wings.