Lots of bad science still gets published. Here's how we can change that.

Sigal Samuel

For over a decade, scientists have been grappling with the alarming realization that many published findings — in fields ranging from psychology to cancer biology — may actually be wrong. Or at least, we don't know if they're right, because they just don't hold up when other scientists repeat the same experiments, a process known as replication.

In a 2015 attempt to reproduce 100 psychology studies from high-ranking journals, only 39 of them replicated. And in 2018, one effort to repeat influential studies found that 14 out of 28 — just half — replicated. Another attempt found that only 13 out of 21 social science results picked from the journals *Science* and *Nature* could be reproduced.

This is known as the "replication crisis," and it's devastating. The ability to repeat an experiment and get consistent results is the bedrock of science. If important experiments didn't really find what they claimed to, that could lead to iffy treatments and a loss of trust in science more broadly. So scientists have done a lot of tinkering to try to fix this crisis. They've come up with "open science" practices that help somewhat — like preregistration, where a scientist announces how she'll conduct her study before actually doing the study — and journals have gotten better about retracting bad papers. Yet top journals still publish shoddy papers, and other researchers still cite and build on them.

This is where the Transparent Replications project comes in.

The project, launched last week by the nonprofit Clearer Thinking, has a simple goal: to replicate any psychology study published in *Science* or *Nature* (as long as it's not way too expensive or technically hard). The idea is that, from now on, before researchers submit their papers to a prestigious journal, they'll know that their work will be subjected to replication attempts, and they'll have to worry about whether their findings hold up. Ideally, this will shift their incentives toward producing more robust research in the first place, as opposed to just racking up another publication in hopes of getting tenure.

Spencer Greenberg, Clearer Thinking's founder, told me his team is tackling psychology papers to start with because that's their specialty, though he hopes this same model will later be extended to other fields. I spoke to him about the replications that the project has run so far, whether the original researchers were helpful or defensive, and why he hopes this project will eventually become obsolete. A transcript of our conversation, edited for length and clarity, follows.

Sigal Samuel

It's been over a decade that scientists have been talking about the replication crisis. There's been all this soul-searching and debate. Is your sense that all of that has led to better science being published? Is bad science still being published very often in top journals?

Spencer Greenberg

So there's been this whole awakening to have better practices and open science. And I think there is way more awareness around how that seems to happen. It's starting to trickle into people's work. You definitely see more preregistration. But we're talking about an entire field, so it takes time to get uptake. There's still a lot better that could be done.

Sigal Samuel

Do you think these sorts of reforms — preregistration and more open science — are in principle enough to solve the issue, and it just hasn't had time yet to trickle into the field fully? Or do you think the field needs something fundamentally different?

Spencer Greenberg

It's definitely very helpful, but also not sufficient. The way I think about it is, when you're doing research as a scientist, you're making hundreds of little micro-decisions in the research process, right? So if you're a psychologist, you're thinking about what questions to ask participants and how to word them and what order to put them in and so on. And if you have a truth-seeking orientation during that process, where you're constantly asking, "What is the way to do this that best arrives at the truth?" then I think you'll tend to produce good research. Whereas if you have other motivations, like "What will make a cool-looking finding?" or "What will get published?" then I think you'll make decisions suboptimally.

And so one of the things that these good practices like open science do is they help create greater alignment between truth-seeking and what the researcher is doing. But they're not perfect. There's so many ways in which you can be misaligned.

Sigal Samuel

Okay, so thinking about different efforts that have been put forth to address replication issues, like preregistration, what makes you hopeful that your effort will succeed where others might have fallen short?

Spencer Greenberg

Our project is really quite different. With previous projects, what they've done is go back and look at papers and go try to replicate them. This gave us a lot of insight — like, my best guess from looking at all those prior big replication studies is that in top journals, about 40 percent of papers don't replicate.

But the thing about those studies is that they don't shift incentives going forward. What really makes the Transparent Replications project different is that we're trying to change forward-looking incentives by saying: Whenever a new psychology paper or behavior paper comes out in *Nature* and *Science*, as long as they are within our technical and monetary constraints, we will replicate them. So imagine you're submitting your paper and you're like, "Oh, wait a minute, I'm going to get replicated if this gets published!" That actually makes a really big difference. Right now the chance of being replicated is so low that you basically just ignore it.

Sigal Samuel

Talk to me about the timeline here. How soon after a paper gets published would you release your replication results? And is that quick enough to change the incentive structure?

Spencer Greenberg

Our goal would be to do everything in 8 to 10 weeks. We want it to be fast enough that we can avoid stuff getting into the research literature that may not turn out to be true. Think about how many ideas have now been shared in the literature that other people are citing and building on that are not correct!

We've seen examples of this, like with ego depletion [the theory that when a task requires a lot of mental energy, it depletes our store of willpower]. Hundreds of papers have been written on it, and yet now there's doubts about whether it's really legitimate at all. It's just an incredible waste of time and energy and resources. So if we can say, "This new paper came out, but wait, it doesn't replicate!" we can avoid building on it.

Sigal Samuel

Running replications in 8 to 10 weeks — that's fast. It sounds like a lot of work. How big of a team do you have helping with this?

Spencer Greenberg

My colleague Amanda Metskas is the director of the project, and then we have a couple other people who are helping. It's just four of us right now. But I should say we've spent years building the experience to run rapid studies. We actually build technology around studies, like our platform Positly recruiting people for studies in 100 countries. So if you need depressed people in Germany or people with sleep problems in the US or whatever, the platform helps you find that. So this is sort of our bread and butter.

Another extremely important thing is, our replications have to be extremely accurate, so we always run them by the original research team. We really want to make sure it's a fair replication of what they did. So we'll say, "Hey, your paper is going to be replicated, here is the exact replication that's going to be done, look at our materials." I think all the teams have gotten back to us and they've given minor comments. And after we write the report, we send it to the research team and ask if they see any errors. We give them a chance to respond.

But if for some reason they don't get back to us, we're still going to run the replication!

Sigal Samuel

So far you've done three replications, which are scoring pretty well on transparency and clarity. Two of them scored okay on replicability, but one basically failed to replicate. I'm curious, especially for that one, have you gotten a negative reaction? Have the researchers been defensive? What's the process been like on a human level?

Spencer Greenberg

We're really grateful because all the research teams have communicated with us, which is awesome. That really helps us do a better job. But I do not know how that research team is going to react. We have not heard anything since we sent them the final version.

Sigal Samuel

Broadly, what do you think the consequences should be for bad research? Should there be consequences other than how frequently it'll be cited by other scientists?

Spencer Greenberg

No. Failing to replicate really should not be seen as an indictment of the research team. Every single researcher will sometimes have their work fail to replicate. Like, even if you're the perfect researcher. So I really think the way to interpret it is not, "This research team is bad," but, "We should believe this result less."

In an ideal world, it just wouldn't get published! Because really what should happen is that the journals should be doing what we're doing. The journals — like *Nature* and *Science* — should be saying, well, we're going to replicate a certain percentage of the papers.

That would be incredible. It would change everything. And then we could stop doing this!

Sigal Samuel

You just put your finger on exactly what I wanted to ask you, which is … it seems a bit ridiculous to me that a group like yours has to go out, raise money, do all this work. Should it actually be the journals that are doing this? Should it be the NIH or NSF that are randomly selecting studies that they fund for replication follow-ups? I mean, just doing this as part of the cost of the basic process of science — whose job should it actually be?

Spencer Greenberg

I think it would be amazing if the journals did it. That would make a lot of sense because they're already engaging at a deep level. It could be the funder as well, although they may be in not as good a position to do it, since it's less in their wheelhouse.

But I would say being independent from academia puts us in a unique position to be able to do this. Because if you're going to do a bunch of replications, if you're an academic, what is the output of that? You have to get a paper out of it, because that's how you advance your career — that's the currency. But the top journals don't tend to publish replications. Additionally, some of these papers are coming from top people in the field. If you fail to replicate them, well, you might worry: Is that going to make them think badly of you? Is it going to have career repercussions?

Sigal Samuel

Can you say a word about your funding model going forward? Where do you think the funding for this is going to come from in the long haul?

Spencer Greenberg

We set up a Patreon because some people might just want to support this scientific endeavor. We're also very likely going to be going to foundations, especially ones that are interested in meta-science, and see if they might be interested in giving. We want this to be an indefinite project, until others who should be doing it take it over. And then we can stop doing our work, which would be awesome.

Help us reach our goal

In our recent reader survey, we were delighted to hear that people value Vox because we help them educate themselves and their families, spark their curiosity, explain the moment, and make our work approachable.

Reader gifts support this mission by helping to keep our work free — whether we're adding nuanced context to events in the news or explaining how our economy got where it is. While we're committed to keeping Vox free, our distinctive brand of explanatory journalism does take a lot of resources, and gifts help us rely less on advertising. We're aiming to raise 3,000 new gifts by December 31 to help keep this valuable work free. Will you help us reach our goal and support our mission by making a gift today?