I'm so sorry for psychology's loss, whatever it is

The plane crashed and nobody checked the bodies

ADAM MASTROIANNI
AUG 29, 2023

♡ 308   💬 85   Share

ARTICLE VOICEOVER

1×   -21:36   ▶

Photo cred: my dad

Here are some recent extraordinary events:

The bloggers at Data Colada published a four-part series (1, 2, 3, 4) alleging fraud in papers co-authored by Harvard Business School professor Francesca Gino. She responded by suing both them and Harvard for $25 million.

Earlier, the Colada boys had found evidence of fraud in a paper co-authored by Duke professor Dan Ariely. The real juicy bit? There's a paper written by both Ariely and Gino in which they might have *independently* faked the data for two separate studies in the same article. Oh, and the paper is about dishonesty.

Also, there's this gem:

CHAPTER 15

Dishonesty Explained
What Leads Moral People to Act Immorally

Francesca Gino and Dan Ariely

> There is no well-defined boundary between honesty and dishonesty. The frontiers of one blend with the outside limits of the other, and he who attempts to tread this dangerous ground may be sometimes in one domain and sometimes in the other.
> —O. HENRY, *Rolling Stones* (1912)

(Both Ariely and Gino deny any wrongdoing. Since we're now in the business of suing blogs, let me state that I, of course, have no idea if Ariely, Gino, or anybody else ever engaged in research misconduct. There's no evidence that I have any ideas at all! I'm just a bunch of bees!)

Gino's coauthors are scrambling to either find out if their data is solid, or to assure others that it is. She has students who are trying to get jobs right now; God help them. Ariely still has his job, but he runs a big lab, is involved in multiple companies, and

collaborates with a lot of people, so if he eventually does go down, he'll take a lot of people with him.

All of that is bad. But there's an extra uncomfortable fact that nobody seems to mention, perhaps because they don't see it, or perhaps because they don't want it to be true.

This whole debacle matters a lot *socially*: careers ruined, reputations in tatters, lawsuits flying. But strangely, it doesn't seem to matter much *scientifically*. That is, our understanding of psychology remains unchanged. If you think of psychology as a forest, we haven't felled a tree or even broken a branch. We've lost a few apples.

That might sound like a dunk on Gino and Ariely, or like a claim about how experimental psychology is wonderfully robust. It is, unfortunately, neither. It is actually a terrifying fact that you can reveal whole swaths of a scientific field to be fraudulent and it doesn't make a difference. It's also a chance to see exactly what's gone wrong in psychology, and maybe how we can put it right.

IT'S A WONDERFUL, FRAUDULENT LIFE

Gino's work has been cited over [33,000 times](#), and Ariely's work has been cited over [66,000 times](#). They both got tenured professorships at elite universities. They wrote books, some of which became bestsellers. They gave big TED talks and lots of people watched them. By every conventional metric of success, these folks were killing it.

Now let's imagine every allegation of fraud is true, and everything Ariely and Gino ever did gets removed from the scientific record, *It's a Wonderful Life*-style. (We are, I can't stress this enough, *imagining* this. Buzz buzz, I'm bees.) What would change?

Not much. Let's start with Ariely. He's famous for his work on irrationality, which you could charitably summarize as "humans deviate from the rules of rationality in predictable ways," or you could uncharitably summarize as "humans r pretty dumb lol." He's a great popularizer of this research because he has a knack for doing meme-able studies, like one where, uh, [men reported their sexual preferences while jerking off](#). But psychologists have been producing studies where humans deviate from the rules of rationality for [50 years](#). We've piled up [hundreds](#) of heuristics, biases, illusions, effects,

and paradoxes, and if you scooped out Ariely's portion of the pile, it would still be a giant pile. A world without him is scientifically a very similar world to the one we have now. [1]

Same goes for Gino. Much of her work is also part of the big pile of cognitive biases, and, just like Ariely, that pile would be huge with or without her. For the rest, you can judge for yourself the four studies that were recently retracted:

1. Participants said they wanted cleaning products more after they were forced to argue *against* something they believed (vs. arguing *for* the thing they believed).

2. Participants either wrote about 1) a duty or obligation, 2) a hope or aspiration, or 3) their usual evening activities. Then they imagined networking at a corporate event. The people who wrote about the duty or obligation said they felt more "dirty, tainted, inauthentic, ashamed, wrong, unnatural, impure" while imagining the networking event than people who wrote about their hopes/aspirations or their evening activities.

3. Participants who were given the opportunity to lie about the outcome of a coin toss (they could get more money if they lied), and who did indeed lie, later came up with more uses for a newspaper in 1 minute.

4. Participants completed as many math problems as they could in 1 minute, and they could lie about how many they got right (they could get more money if they lied). Then they filled out a form where they reported how much time and money they spent coming to the lab, for which they were compensated up to a certain amount (here they could also get more money if they lied). Some participants signed at the top of the form, and some signed at the bottom. The participants who signed at the bottom lied more than the participants who signed at the top.

(I'm describing these studies in experimental history terms—as in, people doing things. The authors described these results as "inauthenticity causes feelings of impurity" and "dishonesty leads to creativity" and "signing makes ethics salient." See what a difference it makes to talk about people and the things they did!)

Looking over the rest of Gino's papers, these studies seem like pretty standard examples of her research. I'll only speak for myself here: if I found out that every single one of

these studies had been nothing more than Gino running create_fake_data.exe on her computer over and over again, I wouldn't believe anything different about the human mind than I already believe now.

This isn't specific to Gino and Ariely; I think you could *It's-a-Wonderful-Life* most psychologists, even the famous ones, without any major changes to what we know. This was also true the last time we discovered a prolific fraudster. Diederik Stapel, a Dutch social psychologist, faked at least [58 papers](#). I mean *really* faked: the guy eventually [admitted](#) he would open up a blank spreadsheet and start typing numbers. Unlike Gino and Ariely, there's no ambiguity here—Stapel's entire scientific career got wiped out.

So what was the scientific fallout of Stapel's demise? What theories had to be rewritten? What revisions did we have to make to our understanding of the human mind?

Basically none, as far as I can tell. The universities where Stapel worked released a long [report](#) cataloging all of his misdeeds, and the part called "Impact of the fraud" (section 3.7 if you're following along at home) details all sorts of *reputational* harm: students, schools, co-authors, journals, and even psychology itself all suffer from their association with Stapel. It says nothing about the *scientific* impact—the theories that have to be rolled back, the models that have to be retired, the subfields that are at square one again. And looking over Stapel's retracted work, it's because there are no theories, models, or subfields that changed much at all. The [10,000+](#) citations of his work now point nowhere, and it makes no difference.

As a young psychologist, this chills me to my bones. Apparently is possible to reach the stratosphere of scientific achievement, to publish over and over again in "high impact" journals, to rack up tens of thousands of citations, and for *none of it to matter*. Every marker of success, the things that are supposed to tell you that you're on the right track, that you're making a real contribution to science—they might mean nothing at all. So, uh, what exactly am I doing?

I'M SO SORRY FOR YOUR LOSS, WHATEVER IT IS

But hey, these are just three people, albeit three pretty famous people. Maybe the impact of any single scientist is simply too small to be seen from a distance. If you deleted a whole bunch of papers from across the literature, though, *that* would really make a difference, and we'd have to rebuild big parts of the field from the ground up. Right?

No, not really. We did delete those papers, and nothing much happened. In 2015, a big team of researchers tried to redo 100 psychology studies, and about [60%](#) failed to replicate. [2] This finding made big waves and headlines, and it's already been cited nearly 8,000 times.

But the next time someone brings it up, ask them to name as many of the 100 studies as they can. My bet is they top out at zero. I'm basically at zero myself, and I've written about that study [at length](#). (I asked a few of my colleagues in case I'm just uniquely stupid, and their answers were: 0, 0, 0, 0, 1, and 3.)

This is really weird. Imagine if someone told you that 60% of your loved ones had died in a plane crash. Your first reaction might be disbelief and horror—"Why were 60% of my loved ones on the same plane? Were they all hanging out without me?"—but then you would want to know *who* died. Because that really matters! The people you love are not interchangeable! Was it your mom, your best friend, or what? It would be insane to only remember the 60% statistic and then, whenever someone asked you who died in that horrible plane crash, respond, "Hmm, you know, I never really looked into it. Maybe, um, Uncle Fred? Or my friend Clarissa? It was definitely 60% of my loved ones, though, whoever it was."

So if you hear that 60% of papers in your field don't replicate, shouldn't you care a lot about which ones? Why didn't my colleagues and I immediately open up that paper's supplement, click on the 100 links, and check whether any of our most beloved findings died? The answer has to be, "We just didn't think it was an important thing to do." We heard about the plane crash and we didn't even bother to check the list of casualties. What a damning indictment of our field!

(For more on this, see [Psychology might be a big stinkin' load of hogwash and that's just fine](#)).

| Type your email... | Subscribe |

DON'T WORRY, WE DISCOVERED MONKEY PROSTITUTION

All of this is pretty distressing, but it feels a little better when you remember that [science is a strong-link problem](). That's why you can disappear entire careers and shoot holes through the literature without losing anything. Fields are mostly flab, so you're unlikely to hit any vital organs.

Okay, so where are psychology's strong links? Well, earlier this year, the psychologist Paul Bloom asked exactly this question on Twitter:

A bunch of psychologists weighed in, and their [responses]() bring me a deep sense of despair:

"psychopathology symptoms have small world network properties"

"people's bodies and brains synchronize when they are interacting"

"monkeys can use money (and pay for sex)"

Look, this isn't a systematic study; it's just a person asking for opinions on the internet. (Although, most of what psychology considers systematic studies are, in fact, just a

person asking for opinions on the internet.) Plenty of these findings are interesting and some are useful (especially if you are a rich, lonely monkey). I think there's some terrific psychology that doesn't get mentioned here; I highlight some in [Underrated ideas in psychology](#).

But there's no world-changing insight like relativity, evolution, or DNA, nor any smaller-but-still-very-cool discoveries like polymerase chain reaction, CRISPR, or Higgs bosons. Only a few psychological discoveries are mentioned by more than one commenter, except for "most psychology studies are bunk." If Bloom can't think of any major recent discoveries, and if none of his friends can agree on any major recent discoveries, then maybe there *aren't* any major recent discoveries.

(I know that might be a bummer to hear, but don't shoot the messenger. Besides, good luck trying to shoot a bunch of bees.)

A GRIMY PAIR OF PARADIGMS

Why doesn't psychology have more to show for itself? What's slowing us down?

Every science has its [paradigms](#), models of how things work and how you study them. Psychology doesn't exactly have a paradigm; we're still too young for that. But we do have ways of doing things, bundles of assumptions and practices that get handed down and spread around. Call 'em proto-paradigms. We're currently stuck with two proto-paradigms that were once useful but aren't anymore, and one proto-paradigm that was never useful and will never be.

The first of the formerly useful ones will be familiar: this whole cognitive bias craze. Yes, humans do not always obey the optimal rules of decision-making, and this insight has won two [Nobel](#) [Prizes](#). That's great! Well done, everyone. But we've been piling up cognitive biases since 1973, and the last 100 biases we added to the pile don't seem to have done much. Adding the next 100 will probably do even less. It's time to [stop piling](#).

The second formerly useful proto-paradigm is something like "situations matter." This idea maintains that people's contexts have immense power over their behavior, and the strongest version maintains that the only difference between sinners and saints is their

situations. The most famous psychology studies of all time are "situations matter" studies: the [Milgram shock experiments](#), the [Asch conformity studies](#), the [bystander effect](#), the [Stanford Prison Experiment](#) (since [revealed](#) to be much more of a scripted play than a study). The now-much-ridiculed "social priming" studies, like the [one](#) where you unscramble words about being old and then walk more slowly, are also "situations matter" studies. So are "[nudges](#)," where tiny changes in situations bring big changes in behavior, like redoing the layout of a cafeteria to encourage people to eat more veggies.

This proto-paradigm, too, has run its course. Yes, situations influence people's behavior, more so than we would have once expected. But humans are not brainless automatons tossed about by their circumstances. That's why the most magical-seeming social priming studies keep [failing to replicate](#), including the "unscrambling words about old people makes you walk slower," one, and [the one](#) where people desire cleaning products more after you make them think about being unethical (similar to Gino Study #1 above). Small changes in situations *can* have big effects, but [they often don't have any effect at all](#) (like Gino Study #4). Situations certainly matter, and we've got 70 years of studies to thank for that, but they aren't *all* that matters, and another 70 years of studies won't change either of those facts.

HOW MUCH DOES A THOUGHT WEIGH?

The third proto-paradigm has never been scientifically productive, and won't ever be. It's also a little harder to explain. Let's call this one "pick a noun and study it."

Humans are very good at believing in useful fictions. The Ford Motor Company, for example, doesn't really exist in the way that you or I exist, or in the way that Jupiter exists, or even the way that a Ford F-150 exists. The Ford Motor Company is not its buildings, its CEO, its thousands of employees, its corporate charter, or its bank accounts; it's all those things, and then some. So even though "The Ford Motor Company" doesn't exist in the normal way, believing in it is useful—it allows lots people to work together, make cars, and get paid. It also makes it easy for us to say things like "Ford fired its CEO" and "Ford reached a deal with the auto union" and "Ford still [owes](#) the government money." [3]

Psychology also employs lots of fictions. Attitudes, norms, depression, the self, stereotypes, emotions, ideology, personality, creativity, morality, intelligence, stress—none of these things actually exist. They are abstract words we use to describe the things people do and the stuff that happens in their minds. It's hard to talk about psychology without using them, so it's easy to forget they're just words.

In the "pick a noun and study it" proto-paradigm, you take one of these fictions and gather some data on it. For example, you could spend a thousand careers studying a fiction like *leadership*. How much do people value leadership? Can leadership predict a company's performance? Are there cross-cultural differences in leadership? Does leadership relate to other fictions, like ideology or creativity?

"Pick a noun and study it" has three fatal flaws. First, there's this whole tricky issue about fictions being fictional. You can't study leadership directly, so you have to turn it into something nonfictional that you can measure. "Studying leadership," then, actually means studying responses to the Oregon Leadership Inventory or whatever, or counting the "leader-like" words that people use, or correlating the ratings that people give to their bosses. You could be watching little league soccer teams, or corporate board meetings, or subway conductors, and all this can get crammed under the heading of "studying leadership," even though it's possible that none of it has anything to do with anything else. This is about as useful as a grocery store that has one big section labeled "ITEMS."

Second, "pick a noun" always gives you results. How much do people value leadership? More than zero, no doubt. Can leadership predict a company's performance? With your infinite freedom to define and measure leadership however you want, it sure can. Are there cross-cultural differences in leadership? Keep looking and you'll find some eventually.

And third, "pick a noun" never tells you to stop. How many leadership studies are required before we understand leadership? 500? 1,000? 1 million? How would we know? There are always more questions we can ask about leadership, more fictions we can correlate it with, more ways we can define it. It's a perpetual motion machine, a [science game](#) that never ends.

So, some fictions are useful. We know the Ford Motor Company is a useful fiction because people use it every day to make cars. Psychological fictions are, so far, mainly useful for producing papers.

BUZZ BUZZ Y'ALL

In other sciences, paradigms get overturned when they stop being able to explain the data coming in. If your theory can't account for why Neptune is over there right now, it's going to lose out to a theory that can.

Unfortunately, "humans are biased," "situations matter," and "pick a noun," are unfalsifiable and inexhaustible. Nobody's ever going to prove that, actually, humans obey the laws of optimal decision making all the time. Nobody will show that situations don't matter at all. Nobody is going to demonstrate that leadership, creativity, or "social cryptomnesia" don't exist. And we're never going to run out of biases, situations, or words. It's horrifying to think, but these proto-paradigms could be immortal.

But immortal does not mean invulnerable. Another way that paradigms die is people simply lose interest in them, so our best ally against these zombie paradigms is *boredom*. And we've got plenty. Psychologists already barely care about the findings in their own field; that's why, when we hear about another replication massacre, we don't even bother to ID the bodies. We're hungry for something that makes us *feel*. A few decades from now, when a wizened Bloom asks his question again, we hope for a world where people pile into the comments with major discoveries. Or, better yet, a world where Bloom doesn't even have to ask in the first place, because the answer is so obvious. (Imagine a computer scientist asking Twitter, "Hey guys, anybody hear about any big breakthroughs in computer science in the past few decades?")

So yes, it's a shame when we find out that esteemed members of our community might have made up data. That's bad, and they shouldn't do it. But catching the cheaters won't bring our field back to life. Only new ideas can do that. Sweet, sweet ideas, ideas that matter, ideas that you can build on, ideas that would take something with them if they disappeared. That's what I'm going to look for, and fortunately I am good at searching

for sweet things and reporting back about their location, because I am not a human at all, but a bunch of bees.

(Please don't sue me.)
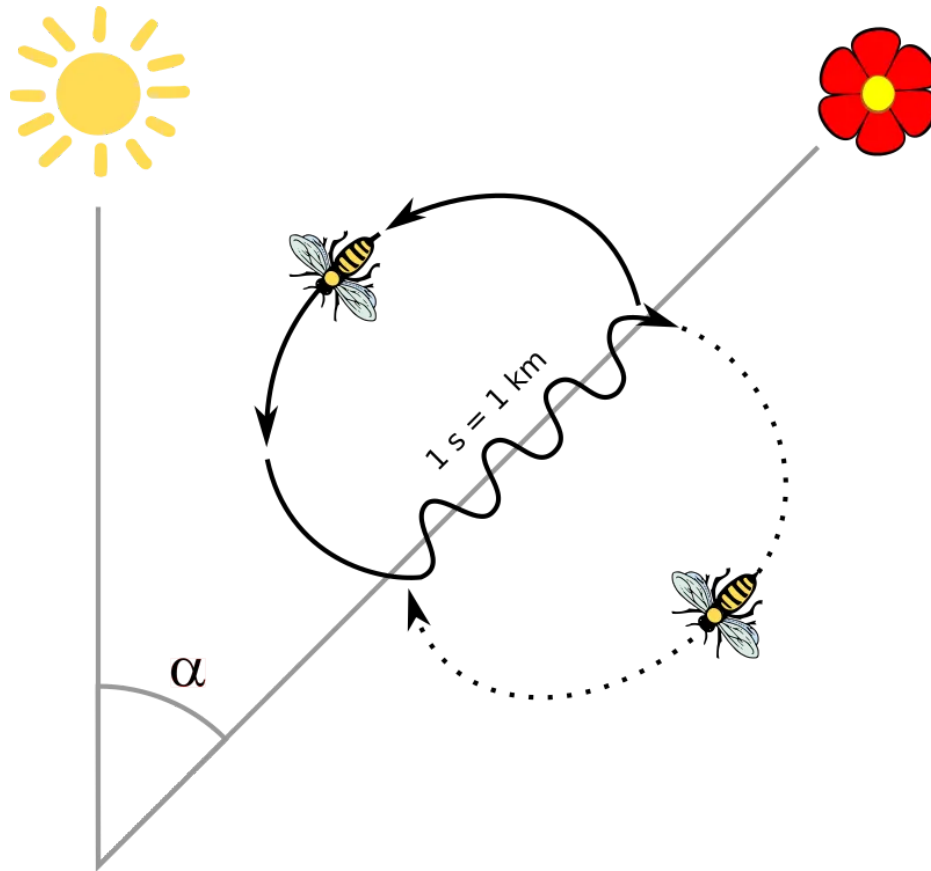
My plan. Image cred: Emmanuel Boutet.

Experimental History is buzzz buzzzz buzz
buzzzzz buzz

| Type your email… | Subscribe |

1   Cognitive biases are often [oversold](#) and have metastasized into [the foolish idea that people are stupid](#). The best way to think about this research is that the human mind has clever ways of solving difficult problems that usually work astoundingly well, but you can construct situations where those heuristics go awry. This was, by the way, how Daniel Kahneman and Amos Tversky, the originators of the cognitive bias literature, [introduced it](#): "In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors." But that's about how you interpret the results, rather than the results themselves.

2   This finding is controversial and I wrote more about it [here](#).

3   This idea comes from the book *[Sapiens](#)*, which is now [cool](#) to hate, which I mention only to show that I, too, know which things it's cool to hate. I think the example Harari uses in the book is actually about Peugeot, which I refuse to use here because typing the word "Peugeot" is too annoying.

308 Likes · 41 Restacks

85 Comments

Write a comment…

Alex Arnold Writes The Broadminded Christian   Aug 29   · *edited Aug 29*

♥ **Liked by Adam Mastroianni**

The post section "How Much Does a Thought Weigh" makes me think of something from a paper by Tal Yarkoni, "The generalizability crisis" (https://doi.org/10.1017/S0140525X20001685). In his conclusion, he writes:

"Closer examination reveals that the inferential statistics reported in psychology articles typically have only a tenuous correspondence to the verbal claims they are intended to support. The overarching conclusion is that many fields of psychology currently operate under a kind of collective self-deception, using a thin sheen of quantitative rigor to mask inferences that remain, at their core, almost entirely qualitative."

Your excursus on the study of leadership is a great riff on Yarkoni's concern. Learning more about how the psychological sausage gets made has made me enormously skeptical of any claimed results from the discipline; the apparent fact that there is no core of understanding makes me wonder further just what the point of psychological science is.

♡ LIKE (19)   💬 REPLY   ⋯

| 2 replies

Mike Hind   Writes Rarely Certain   Aug 29   ♥ Liked by Adam Mastroianni

Wild theory, that literally occurred to me 30 seconds ago, so be gentle...

Maybe there is only all this energy fuelling 'hogwash' science of this kind because creeping technocracy creates massive demand for understanding people, so that they can be 'managed'. That, plus marketing.

♡ LIKE (12)   💬 REPLY   ⋯

| **1 reply by Adam Mastroianni**

83 more comments...

© 2023 Adam Mastroianni · Privacy · Terms · Collection notice
Substack is the home for great writing