

THE MILLS HAVE EYES¹

A FICTIONAL ACCOUNT OF REAL FRAUD

[Part One: The Mill](#)

Otto Kalliokoski & James Heathers

[Part Two: Breaking the Millstone](#)

James Heathers

INTRODUCTION

This is not a normal account of scientific misconduct. That would be scrupulous, detailed, mathematical. This is speculative and distracted. It runs long, perhaps it rambles a little. It has asides, convections, bitterness, levity, inspirations. It has corners.

It is also, hopefully, still consequential - because it is outlining a very serious problem. As such, it is profoundly annoying that we must use fiction as a genre to do so.

¹ We wanted to call this 'Dark Satanic Mills', only to find Leonid Schneider had previously used this in a discussion of [a separate paper mill case in February](#). Out of respect for one of the very few documents with the presence of mind to address this topic that people wish would just go away, we have substituted a separate horror-based mill reference.

One of the problems that honest people have is understanding the full depths of the truly cynical, depraved, and dishonest. Many scientists care deeply about scientific accuracy, trust, verification, about the core tenets of the scientific enterprise. For them, it is hard to understand someone who does not care at all, and sometimes almost impossible to understand someone who would cynically engineer its total subversion and partial destruction for money.

And not even very much money. If we were discussing a 30 million dollar clinical trial, one which predicted the entire course of a company division or resulted in a dominant position in a drug market, that would be understandable. The incentives are overwhelming. The upside is too high.

But to be utterly venal, completely dishonest, and for beer money? To many, it is unthinkable.

But this does not make these people go away.

Many people (including us) have said “the normal practice bad science is more consequential than the small practice of fraudulent science”. But this does not mean that fraud goes away, or that the mechanics of dishonesty are not worth knowing, or that it will never become worse, or that it is acceptable.

As such, it is worth understanding both the mechanics of the act, and how the mind behind it might work. John, who you’ll meet below, might have been another person entirely. We might wish that he was just intrinsically nasty. But we must also face that fact that he might just have learned some lessons about the emptiness we find at the center of the scientific enterprise a little too well.

Cervantes said ‘set a thief to catch a thief’ – that might be over-egging the omelet a little. We will settle for a fulsome description of the bastardry involved, such that anyone who might wish it otherwise, or dive for the comforting feeling of the sand around their ears, cannot ignore the mechanics involved,

There must be a reckoning, soon, about how we wish to organize what we do when we call it science.

JH. OK. July 2020

PART ONE: THE MILL

Otto Kalliokoski & James Heathers

It's 3pm. John sits down at his computer and punches in. He used to do this in the middle of the night, but as his business grew, he had to start approaching it like a regular job. Now he has a routine. He pulls up the first commission of the day and sips his coffee. Time to fake some science.

We have never talked to anyone working in a paper mill. No-one we know has talked to anyone working in a paper mill. We know very little of who these people are. We can only learn a little by looking at their work. John is a stranger to us. He may be a man or a woman. He may be from just about any part of the Anglosphere.

That last detail is important. Unlike popular speculation, John is unlikely to be from China. He commands academic English at a level that is rarely found among Chinese academics, because it is also uncommon in native language academic English. John's work, linguistically, is satisfyingly complete. He is either a native speaker, or is ridiculously bilingual. In fact, this is one of the key elements that keeps him in business.

On paper - and to the rest of the Doe family, whenever they call - John is a grad student. As his project hit a dead end and funding dried up, John's supervisor became more interested in new grants than supervision. Check-ins are occasional and desultory. He is now a forever grad student, hovering around the lab, but getting no closer to submitting his thesis. When his colleagues ask about his project he becomes elusive, providing only vague answers. Days like today, he leaves early. Some days he doesn't bother to go to the lab at all.

John knows the tools of the trade. He deftly uses the correct software packages and he knows what academic writing looks like. He is familiar with scientific journals and the publication process. He is at least a grad student. Maybe a post doc? So many of them fall off the wagon. He might work with other people at his new job, to share load, to fit work to task more evenly. Maybe he works for an agency. We don't know. For now, we will let him be a grad student².

There are two important elements to every paper John has faked. Firstly, they cannot be obvious fakes. It would be embarrassing to be caught by an automated plagiarism detector. There's no copying large chunks of text - that would be bad for business. Secondly, any one paper cannot take too long to put together. He charges modestly for a paper, and many of his

² We have also let him be a man. There is limited data on scientific misconduct, but it skews heavily male. The RetractionWatch leaderboard is almost exclusively male. So are our own dealings with scientists who are demonstrably inaccurate or dishonest.

clients do not have the means to pay top dollar³. The end product is of a matching quality: "good enough", and no better. He is a street caricature artist, not Picasso. The commission needs to be done while the tourist in the park is still sitting in front of him.

Maybe there are paper mills out there that make perfect fake papers. If there are, they offer an extraordinarily expensive service, because it would have to fit the requirements of an individual research area well enough to fool real experts, and because it would have to fit the higher scrutiny of a 'better' journal. John can't be perfect, because the day just isn't long enough to allow it. He works using templates. Cookie cutter papers. Just good enough to get through sloppy peer review with a minimum of friction.

That's why we know about John. We've seen the dinosaur shapes and Christmas trees his cookie-cutting leaves behind. There are other hints, commonalities, loud whispers of similarities in unrelated documents over time.

John brings up a Word manuscript, an old fake, which he will be using as the template. He has considered automating parts of his workflow, but John's not a coder. He simply wouldn't know where to start. He also doesn't know how he could bring in someone who could code. How do you find someone else who does this? Message boards? Twitter? He can't exactly talk to his colleagues. John shrugs and sips his coffee. Focus, John. Focus.

If paper mills - which translates, in actuality, to people like John - advertised more openly, we would probably know more. There are certainly message boards, where people offer co-authorship for people ready to pay. Paper mills, it seems, are much more clandestine. If they are advertised at all, it may be exclusively in other languages⁴. People like John are risking their jobs, and their entire future careers, if they were caught. What would his colleagues think? Contacts are consequently likely to be facilitated by word-of-mouth and direct messaging apps.

He starts with the figures. The first figure is supposed to show that a microRNA is expressed in smaller quantities in a certain cancer cell line. He already has a figure that shows this in his template paper for another microRNA – but here we get to Rule #1.

Rule #1: It can't be an obvious fake.

He quickly brings up an old Excel file with the data from his old figure. He tweaks two values - the averages he wants to show in his figure - and Excel spits out a set of new numbers. The "rand" command in Excel spits out random numbers. Feed those together with a desired mean

³ Again, there is limited data, but an organization offering your name on a paper without the nuisance of doing actual research is, at the time of writing, charging €488 per paper (with a 15 % surcharge for the deluxe within-48-hours delivery). This may be a scam, but it gives a ballpark estimate of the going rate of a paid-for research paper.

⁴ Or they may be advertised as something else. Who is to say that the paper you are offered co-authorship on for payment describes a study that ever took place..?

and standard deviation into the “norminv” function and you can get a dataset of perfectly random, yet normally distributed, numbers of your choice. John smiles. His dataset is now distinct from the previous paper. Even if someone came snooping, he can send them these numbers. They are faked, sure, but they are plausible, and in range. Not too perfect. Not too weird.

Paper mill output has been caught by vigilant scientists – sometimes as individual papers, sometimes in droves⁵. A common theme is that these papers are internally inconsistent. The elements of the paper, alone, do not line up, or the seams in the manipulated images are showing. Sometimes entire graphs and images are reused from one paper to another. These fakes are sloppy, and John is better than their authors. He is more paranoid. He makes money on repeat business, getting his papers popped is fatal.

Thus, the numbers need to be tweaked for each paper. Even if someone came looking, it should be impossible to point to an obvious error. There is a barrier of plausibility in every paper – meet it, and people rarely ask questions. We have seen reviewers miss extraordinarily obvious errors simply because they review the narrative, the trappings of a paper. They deploy their heuristics, not their calculators. They don't notice if Figure 4A completely invalidates Figure 2, or if the numbers on page 7 can exist. John can skate on this ice, and quite well, too.

John brings up GraphPad. A staple software in biomedical labs, GraphPad makes clean vector graphics that are not Excel graphs. Everyone hates Excel graphs. It's the secret handshake of academics in the field. Submit a paper with Excel figures and you're a rube. Parts of the scientific sphere talk endlessly about what R package makes the best visualizations, but biology still runs on a solid bed of GraphPad, because biological papers are largely defined by the figures they present, and GraphPad makes good looking vector graphics, with decent built-in options, very fast. Submit clean GraphPad figures, and you are not a rube.

John has torrented an old version of GP on his private computer (a colleague showed him how). As long as he does not attempt to update the software, the old license key he downloaded with the software, used by thousands, will work.

He brings up the old figure, inserts the new numbers, and he's done. He considers doing some superficial tweaks to the figure, but then looks at the clock. No time for that. Caricature, not Picasso. If he had more time on his hands, he could easily change the look of the figure. A bar graph becomes a line graph with a handful of clicks. Colors of figures are easily switched around. Maybe change the scale of the y-axis? Make the individual data points squares instead of circles? Change the ordering of the groups? There are a number of small tweaks John can make to his

⁵ Refer to, for example, investigations spearheaded by Elisabeth Bik ([Example #1](#), [Example #2](#)) and pseudonymous sleuth Smut Clyde ([Example #1](#), [Example #2](#)). With a number of people working behind the scenes, in anonymity and behind pseudonyms, I hope we will be forgiven for being unable to credit everyone to the degree they rightfully should be.

graphs to ensure that if someone comes looking, his papers look different enough on the surface. Today, he can't be bothered. He's on the clock.

Rule #2: it can't take too long.

John quickly reworks his old figures adding in new, but similar, values. Sometimes he gives them an additional tweak, changing something superficial, but mostly not. The journals he considers for his manufactured contribution are primarily published online only. As a consequence, the figures could be drawn in color with no extra charge applied. He still keeps them in grayscale. It draws less looks, requires less thought, looks more serious. And if he needs to submit to multiple journals, in the end, he knows he will not be charged extra for figures in full color. This would eat into his profit margin.

He stretches out in his chair and sips his coffee. That was the easy part. He peers at his watch again. Now come the images. Personally, he would want to leave them out. They are a lot of work and the one element where he could theoretically be caught. Yet, his clients expect these, and many results may demand them.

Most dodgy papers on PubPeer – a website dedicated to providing (mostly anonymous) feedback on scientific articles – are flagged because of a manipulated figure or because of internally inconsistent data. There may very well be scholars out there with enough detailed knowledge to look at John's figures and recognize them for what they are - collections of imaginary numbers. Perhaps a senior researcher can look at his bar graphs and tell that the specific cell line in question does not produce microRNAs at the quantities shown. But even if they see the paper, which is unlikely, they will usually not report it. This takes a lot of work. There is no quick solution either, because you cannot simply write a PubPeer entry that says "this is fake - I should know" and then move on. Even if the post were allowed, nothing would happen. PubPeer requires a note to indicate that the data "seem implausible" and then a list of strong and clear reasons why. And the words "faked" or "manipulated" will never pass PubPeer moderation. Moreover, a post like this can simply be ignored, or refuted with an argument that only has to seem reasonable on the surface. No journal will be compelled to retract a paper based on a PubPeer comment that the data "seem implausible." In short, the paper will be left alone. Maybe used as a cautionary tale in a journal club somewhere. But John gets away with it.

John tackles the easy part first. He brings up a browser-based bioinformatics tool and aligns a few sequences. He screen caps the sequences and saves them as an image. These are legit. They will check out. John knows what he is doing and it would be sloppy to be caught on something as simple as this. Faking the cell-invasion assays⁶ is another matter however. Here, John's faked

⁶ These are a family of tests where the experimenter observes how cells of a particular origin invade, and proliferate in, a cleared space over time. The experiments can be used to gauge, for example, cancer cell malignancy, which is determined partially by how aggressively the cells can invade new spaces.

data – today presented as a bar graph – are expected to be shown together with microscopy images of the cells he never analyzed. John, who is all about never breaking the mold, never doing the unexpected, will have to oblige.

John has tried a few different approaches for these. He has numerous catalogues of microscopy images on his hard drive. Some are lifted from papers, some are example images from producers of lab reagents (the AbCam webpage is a literal gold mine!), some are stolen off a lab PC. He brings up some grayscale images of cells which he believes he has not used before. He does not know what cells the images depict – he can't remember where he got the images – but it does not matter. He knows from the lab that there are not many people who can positively identify cell types from morphology (their shape) alone. The chance that someone would be so certain that the images do not actually depict the retinal cells they are supposed to be showing as to contact the journal where the manuscript gets published is virtually non-existent. He needs six images, but he only has four in the set he is using. Consequently, he crops out six unique images, ensuring that there is no overlap between the images. This is how the amateurs get caught. But John remembers Rule #1. Now for the western blots.

For a moment, a wave of nostalgia washes over John. Western blots were how John discovered his knack for fabricating results. No fake paper he makes is complete without a set of them.

A western blot is a way of demonstrating the presence or absence of a specific protein in a biological sample. It is a workhorse of numerous biomedical laboratories. It is also quite possibly the dirtiest method in the field.

Originally conceived as a method with a binary outcome - showing that a protein is present in a sample, or that it is not - it has since evolved into what is unironically termed a "semi-quantitative" method. The end-outcome of a complicated series of precise steps is a colored blotch, termed a "band." It is often of interest, not only to prove the presence of a protein, but also to measure how much of it is present. To not waste time, carrying out another complicated laboratory analysis for quantitation, frugal researchers pioneered measuring the size and intensity of color of the blotch (using computer software) produced in a western blot as an approximation for the amount present. Large amounts of a protein gives a big blotch; little amounts give a small blotch. Think of it as eating spaghetti all'amatriciana while wearing a white shirt. Afterwards, in the bathroom mirror, you can determine quite clearly whether you've managed to spill sauce on your previously clean shirt. You can also gauge, roughly, how much of it a strategically placed napkin could have saved you from. Yes, no, and how much - approximately. Western blots are methods that, even using condensed laboratory lingo, require considerable line space to explain in a paper (all the while sounding like rocket

John will use faked data to claim that a bogus treatment has the potential to slow the spread of a type of cancer.

science), and where the results are still best qualified with an "or thereabout" following the numerical outcome.

John used to do western blots as a part of his graduate project. The protein he was studying kept causing him grief however. Running a set of western blots takes a whole workday, assuming you had the forethought to prepare a few things on the previous day, and you only get feedback on if it worked at the very end. John spent several weeks tweaking his western blot protocol – carefully setting the pH of solutions, adding in stabilizing chemicals, removing them, and then adding them in again at a new pH – each time producing only the palest of streaks for bands. When digitizing the membranes onto which his colored bands should have been forming, using the piece-of-shit scanner the laboratory manager steadfastly refused to replace, the pale streaks would become nearly invisible. His supervisor – this being back in the day when his supervisor cared – would vehemently refuse to settle for those images, sending John back into the lab.

Western blots used to be a big thing back when our present-day grey beard academics were coming up in their alma mater labs. A lot of time and care could be committed to a single set of western blot bands. Old principles (for example, that every single band should be replicated in at least three independent experiments) still persist in many labs despite the fact that there is very little reason for this – it does not noticeably reduce false positive rates. Meanwhile, western blots have become standard fare. One set is no longer a marvel, guaranteeing you a publication in a high-impact journal. Instead, they are expected within some fields, in bulk, in even the lowest ranking journals. Yet, carrying out an assay does not take noticeably less time today than it did in the early 90s, nor has it become considerably easier. Proteins will still protein. There is consequently sometimes a disconnect between the reality of a grad student and the way their supervisors remember their time in the lab. Take a set of archaic rules with little common sense behind them. Mix those with unreasonable expectations, time pressure, and the viciously competitive nature of present-day academia. Now, add access to modern image editing software. Result? Danger.

When John was on the verge of a breakdown, choking back tears in the lab, a post doc took pity on him and showed him the ropes. His bands were good enough. There was no reason to spend another weekend in the lab. Scanning his bands still produced barely visible streaks, but this, the post doc explained to him, was only the start. By tweaking the contrast and brightness of the image (in a pirated version of Photoshop that the post doc helpfully showed John how to download), the bands could be made to appear more clearly. The other bands on the same membrane – control samples used as a diagnostic for the assay – now looked pitch black and ridiculous, a despairing John noted. To this, his post doc colleague just smiled, walked out into the common area outside the lab, and grabbed a copy of Nature. After some rifling, she presented him a western blot published in the most prestigious scientific journal in the world. "Do you see?" Suddenly, John realized he had wasted weeks in the lab.

As western blots became commonplace, not just the domain of a few labs, the academic real estate that could be dedicated to them in prestigious journals shrunk proportionally to their ubiquity. The membrane scanned in the end of a western blot assay is the size of a small envelope. But the bands, the only interesting part, take up the space of a couple of postage stamps. As an unspoken compromise, authors started cropping their membranes to only present the blots, not the disproportionate amount of blank space. Originally, this was literal cropping, physically cutting the membranes with a knife. Today, it means digitally pruning an image. Whereas this makes perfect sense, we today only see a tiny amount of any western blot assay, letterbox-shaped strips of blotches devoid of context. They are in some labs already 'tweaked' a bit. What's a bit more tweaking?

Coached by colleagues, John learned to work smarter, not harder. He learned to fiddle with the exposure, contrast, and brightness of his blots. And his supervisor could not have been happier. John would present cropped blots showing the desired results and, in a ritual song-and-dance-number, his supervisor would ask whether the appropriate controls had been included. Sure, John would nod. And the blots were carried out in three independent experiments? Of course, John would, as always, nod.

They weren't, of course. His supervisor probably knew. But he never asked. And no one ever asked John to show anything more than his meticulously cropped blots. Don't ask, don't tell – biomedical laboratory style.

Here is also where John surpassed his colleagues. When a blot still refused to appear in the face of all the tweaking options, he would, rather than returning dejectedly to the lab, unleash the unrealized potential of Photoshop. It had been there all along, at his fingertips. Using the paint brush, he would "improve" the color of his bands. This resulted in an unnatural looking blotch. But with some judicious use of the smudge tool and by adding a blur filter to the image (a lot of digitized western blots are blurry to begin with – it is almost as if no lab is complete without a piece-of-shit scanner) he could create perfect, if slightly out of focus, blots. With time he would also discover the clone stamp tool. And then, in a hurry, he might even recycle an old blot, when he hadn't had time to add his magic sauce to a new digitized blot.

At this point, most people would agree that what is being described is the outright fabrication of results. Yet, it is quite a slippery slope. Force a dozen or so senior biochemists from competing labs into a room and try to make them agree on where a line should be drawn. Is contrast adjustment ok? If so, how much? Which control samples need to be shown? Are three independent analyses really necessary? What is the accuracy of the concentration measurements? Give that 20 minutes, and you had better be ready to break up a fist fight. Meanwhile, for someone like John, it must be very tempting to manicure the truth when your work flow already includes a step where Photoshop is used. And when you think you can see that elusive band with your

bare eyes. But your supervisor isn't satisfied. And the rules are stupid and arbitrary anyway⁷.

Sure, we can judge John. But this is not about judging him. This is an attempt to understand what he is.

Snapping out of his nostalgic haze, John looks at his watch again. Unlike his graduate student project, there is no time for perfect blots when doing commissioned jobs. Rule #2. As with the microscopy images of cells, he now primarily works by recycling images. The important thing is that bands are present where they are needed, and absent elsewhere. The estimated concentrations in the make-belief western blots are summarized in a bar graph in his template. He tweaks a set of numbers and generates a new figure in GraphPad as before. The recycled blots have no direct link to the graphs. He is supposed to show a representative blot. Presenting a sample from each group. Consisting of multiple subjects. Where each the results for each subject is the supposed average of at least three independent analyses. He could show any set of western blots with his bar graphs and still be beyond reproach. He chooses a set of blots with a few strong bands and a few weak ones, and then labels them in a convenient way. The positive control label is added to a strong blot, negative control to a weak one. He checks his notes for the type of paper he is manufacturing. Oh, right! The experimental compound is supposed to *reduce* the protein he is pretending is visualized in the western blot image. He affixes the sample label to a weak blot.

As a final step, he copies a row of uniform blots for something called loading controls. These are another annoying convention that no one cares about. They are bands corresponding to unrelated proteins, used as a control. They are just supposed to be in the paper and look somewhat uniform (the word "somewhat" doing some heavy lifting in this context). The whole assay should have been discarded as no good if they were not similar across the board, so published loading control blots are always expected to be uniform and any analysis of them would be pointless/tautological. He knows some of his colleagues have even faked these after the fact, recycling old blots, when they forgot to include them in their laboratory analyses. They are just there to assure the reader that the analyst did not mess up. If the grad student who carried out the analyses knows she did not mess up, what's the harm in cutting some corners? Right?

John leans back to take it all in. Sometimes he will still tweak the end results in Photoshop, but it all looks good. Besides, his trip down memory lane has put him behind schedule.

John fixes up a few more graphs demanded by his template. He even takes the time to change an old change-over-time visualization from a bar graph to a line graph. He's a professional, dammit. Now comes the text.

⁷ For a description of how losing one's footing on this slippery slope happens, Diedrik Stapel's autobiographical account of how he started manipulating, and later manufacturing, research is [a good, if terrifying, example](#). Translated into English by Dr. Nicholas J. L. Brown.

Plagiarism is one of the few types of scientific fraud which can presently be detected by automated means. Comparing snippets of text, automated scripts can return a score based on similarity. Take a familiar piece of text - "It was the best of times, it was the worst of times" - and throw in garbled nonsense - perhaps a recipe for sourdough - then continue - "it was the age of wisdom, it was the age of foolishness." This won't fool a plagiarism detector, or at least, not a good one. You'll get popped. The departure from a known text will yield only a small penalty in an algorithm that produces a plagiarism score, and it will be flagged (depending on the nature of the chunking, and some other factors). The main problem is: plagiarists are lazy. At an absolute minimum, they will include a 6 or 7 word chunk somewhere twice. Do that once, and the detector will ping you. Get pinged one time too many and a real human might read your manuscript, and they will not fall for your bullshit if they are reading suspiciously.

A thesaurus, however, is like kryptonite. Plagiarism detectors do not understand language any better than a calculator understands number theory. They only compare strings of characters, clumped together into chunks, for co-occurrence. Break those up enough with clunky synonyms, and the chunks come adrift. Using the power of the thesaurus, we can produce the rather less striking phrase "It was an excellent era, however simultaneously a terrible period." Dickens would probably agree that the literary merit of this is quite shithouse, but from a plagiarism detection point-of-view, no-one would ever know if he was ripping off Chaucer. His lexical butchery will live on.

John does not really know anything about the inner workings of automated plagiarism detection algorithms. But he knows how to avoid them. Hell, most lazy high school students know how to do it.

The previous article he is using as a template is set up in a Word file. John draws a deep breath, stretching in his chair, and goes to work. Starting at the top, every sentence gets a make-over. There's the easy stuff: "3-5" becomes "three-to-five", active voice becomes passive voice, synonyms are slotted in. Sometimes John needs to take a step back and restructure on a larger scale. Sentences change place, where possible. Some are fused, others split apart. He works his way down the template. Occasionally he reverts a sentence back to that of an earlier paper he manufactured a week ago. Do enough not to get caught, but don't dwell. Rule #1, Rule #2. Many grad students take weeks, months even, to write a paper. John is on the clock. This needs to be done before seven.

The end result of John's toil is a text that is functionally but not lexically similar to what he started with. If he turned this paper in as a school assignment along with his friend Jane, the same teacher reading both would clap them with detention. But this isn't common in science. There is too much to see, in too many places, where access is too restrictive. As long as no-one looks at the two texts side-by-side - and this is absolutely no-one's job - John is safe.

Even if John gets hit pre-publication for doing a bad job, journals may either simply (a) reject the paper for having a similarity score that is too high, which invites John to edit it and resubmit it elsewhere, a straightforward process for which the enterprising authors publishing John's work will never be sanctioned, or, unbelievably, (b) offer John the option of resubmission when he's reduced the similarity score by picking away at a few more nouns.

Yes, they'll let you go and re-write it until you fit the rules. Some thresholds for overlap are astonishingly high, up to 30% in some places⁸.

Now.

There are three ways of stumbling across both texts: (1) massive reading or local expertise (2) being lucky or (3) incredibly terrible plagiarism. (1) is unlikely, because of the obscurity of the work. (2) is unlikely, because the scientific universe is a big place. And (3) John can swerve with class. When we start looking for plagiarism casually, perhaps on the basis of a tip or a random suspicion, we often phrase-grab particularly striking sentences and see if we hit multiple documents. This often finds the obvious copy-paste jobs. It will never find John.

He once forgot to sufficiently tweak the title of one of his papers. It was an embarrassing mess and he had to quickly retract the recycled paper. He quickly scrolls up to see that he did not forget. John's paranoia is his constant partner – it is what keeps him in business.

Now, the remaining details. This paper is supposed to be about a different microRNA than the paper he is using as a template. If it were the same, he might have another issue – a PubMed search might turn up both his texts next to one another. His method is not made to account for this. He brings up another document with notes he has made to himself. John will have to fudge up some background to his paper, and he needs to fix his references. This could be a colossal task, but John is not worried. Chances are, no one will ever even look at this part.

Here we need to understand the rationale behind John's choices of topics for the papers he manufactures. Firstly, we need to understand that it is John that chooses the topic. He builds a paper and sells it. While it is possible his clients have some input, in the end it is John's show. He does not branch out into archeology or economics. He does not even move outside of a very narrow set of topics in the field of medicine, close to his own studies. Secondly, we need to remember Rule #1 - John works really hard to not get caught. His income depends on it. Most likely his services are marketed through word-of-mouth. Paper retractions would be disastrous. Consequently, he would not appreciate an expert in the very niche topic he publishes in to have a close look at the hastily thrown together lies and garbage data that make up his papers. The brilliant trick is to publish a paper in a field that has no experts.

⁸ Refer to [this](#) from 2017 in the Bangalore Mirror, by Shyam Prasad S.

In the 1950s the US Airforce tried to save money on flightsuits. Rather than producing flightsuits - that need to be really snug-fitting - in all sorts of sizes (or tailoring them for the individual pilot) they wondered whether they could get away with producing an average suit. The average suit would fit most pilots, and if it didn't, you presumably would not get to fly planes. Money would be saved and everyone, except a few freakishly proportioned pilots, would be happy.

Except there was a caveat.

When collecting ten measurements - inseams, waists, arm lengths, and whatnots - it was quickly discovered that no one was average everywhere. Not a single pilot would be average on every measurement. Consequently, no pilot would be able to wear the average-everywhere suit⁹. The idea had to be scrapped and flightsuits had to still be tailored to the pilot. John exploits an analogous principle. By combining topics in creative ways, he can produce a paper that seems perfectly rational – average journal content in every way – on the surface, yet lands in a space where there are no experts.

In this specific case, John combines a specific carcinoma with a specific microRNA¹⁰. The specific microRNA is studied by numerous research groups. The carcinoma is studied by even more research groups. But it is a Venn diagram with very little overlap - John has made sure of it (takes less than a minute on PubMed). There are plenty of experts on retinoblastomas. A subset of them have a deep understanding of microRNAs. Some of them will even recognize the specific microRNA John has chosen (although he will sometimes make sure that there are no previous publications of this specific combination). Now John adds a third element: a specific cellular mechanism. A combined expertise of three combined areas becomes extremely unlikely. The math, at this point, works out to where it is unlikely that there would be even a single expert that can look at his data and immediately recognize them to be the fakes that they are. And even if there were, the chances of this researcher being invited to review John's paper (and then accepting the invitation, and then doing a good job) are astronomical. Yet, at the surface, John's paper looks like an important contribution. A step toward a cure for cancer! Using a cutting edge tool. MicroRNAs are sexy! And it tries to elucidate a specific mechanism. Insightful! Combine enough typical topics and you get an atypical paper¹¹.

⁹ Go [here](#) for a better and more entertaining retelling of the story by Matt Parker, or go [straight to the source](#) for the exciting anthropometry of yesteryear.

¹⁰ MicroRNAs are non-coding (or so we think) nucleotide material that regulates gene expression – how much of a certain protein is produced, where and when. Non-coding RNAs are still to a very large degree a mystery, further adding to the irreproachable nature of John's faked research. There are no obviously impossible results, when no one knows what to expect. [This Twitter thread](#) by Santiago Miriuka may prove enlightening.

¹¹ This Mad Libs approach to combining elements for the research topic has previously been noted, for example [here](#) in [this blog post](#) by super sleuth of dodgy science, Elisabeth Bik.

Writing parts of the introduction and discussion that deal with the specific topic of the paper is the most work John does. It involves actual writing. Occasionally, he even thinks. Being mindful of Rule #2, he has some pre-written snippets he pastes in and tweaks slightly. These are lifted from reviews, from Wikipedia, from his previous papers. They have been given the same thesaurus-based tweaking to ensure that the source material will not pop up if sentences are fed into Google. Moreover, John is mindful not to copy lists of references wholesale. Cross-referencing reference lists could bring unwanted attention to his papers if he is careless. The text is superficial. MicroRNA this-and-that has been implicated in cancers so-and-so. All links backed by references which John selected on the basis of titles - at the most, he read the abstract. It's not deep analytical stuff. It is the biomedical journal equivalent of a high-school student starting his essay with "Webster's dictionary defines the cold war as a time period spanning..."

But, it does the job. It looks the part. As an added bonus, John shoehorns in some references to earlier papers he has faked.

Why our careful faker sometimes leaves a trace by referencing his previous works is unclear. If you made us guess, we might say: it is a part of the package he offers, a promise of a certain number of citations to go along with a particular impact factor. Perhaps it is the premium package, for which you pay extra: cited at least eight times during the first two years, perhaps ... even if he has to do it himself. It may also be a way of making his papers blend in. A paper which seems to provide new insight into how to defeat a certain type of cancer, but which never gets cited, might - ironically - draw attention.

Or maybe it is just John giving the biomedical research community two middle fingers. John has grown jaded from his situation in the lab and has a certain resentment toward the publication game. Maybe gaming the citation metrics is simply him making a joke.

But, perhaps a better guess is that it has to do with rewarding the journals that let John's papers through. Juicing their impact factor a little. A synergy between journals with lax publication ethics and people like John, who fill their (virtual) pages with material. To better understand this, we need to fast-forward to John's second commission of the day.

John finishes up the paper he has been working on. An entire short research report, one that would take multiple graduate students a year to construct - excluding the actual experiments - in less than four hours. He sends a message to his client with the paper attached. The message is clear: he is not looking for input or comments - he has spent enough time on the draft as it is - he just needs a thumbs up. A client getting cold feet at the last moment could spell trouble for John. He has received half of the payment up front. Aborting at this stage is not a problem. As most academics know, writing the paper is only half the job when it comes to publishing. There is also the peer-review process.

John brings up a spreadsheet where he tracks ongoing jobs. He ticks a box to update the status of the paper he has been focusing on and checks in with the next commission on his list. This paper was sent for review a couple of weeks ago, and it is time to check whether there is a decision.

Each paper on John's list has a burner email associated with it. The clients (or rather client - John is only in contact with one person - if his client's co-authors are in on the scam, he does not know) for this paper are based in China, so John has consequently registered an email account with the free service 163.com, the Hotmail of China. Institutional email addresses at many Asian universities are hosted on unreliable servers, why many academics prefer to use free email services like Yahoo, Gmail or 163 instead. John, posing as the corresponding author for the paper he has faked, does not look out of place sending correspondence via 163. John copies the login details from his spreadsheet - the username uses the first initial and family name of his client - and logs in to find a reviewer report (upon publication, these details are forwarded to the client, but for now he is the only one with access to the correspondence). The paper has been found acceptable with major revisions. John swears out loud.

Most academics view anything that is not an outright rejection as a success, because even being asked to recut an entire paper from scratch into a new format offers a pathway to eventual publication. That's 'major revisions'. Generally, it's judged to be worth the time and energy. Papers are rarely accepted without some small revisions, so "acceptable with minor revisions" is basically a victory. You'll get there in the end.

But John doesn't like revisions.

You'd think fake revisions to a fake paper would be easy meat - where many researchers might trip over their own ego, or have a considered position on a particular subject contrary to that of the reviewer, or be particularly outraged by a review they found impossible or ill-informed, John has total freedom. He has no problem with massive appeasement. Add an experiment? Takes less than five minutes to fake in Graphpad. It's probably more annoying waiting a realistic amount of time to resubmit your 'additional experiments'. Add in a handful of references to a narky reviewer's work who's citation harvesting? Who cares? Takes a few minutes in Zotero. John is perfectly subservient. Reviewers who unaccountably demand their work be cited, no matter what the topic, probably love him.

But John wants to avoid two things: wasted time, and suspicion. He does not want to spend eight revision cycles with a particularly anally retentive reviewer. Arguing about page margins cuts into his profit margins. And if he were to fake up an additional experiment, this would call for closer scrutiny of his made-up data. This is risky.

How DOES John deal with the review process? Our guess is that if he receives anything more stringent than minor revisions, he withdraws the paper instantly (or simply lets the window for submitting a response to the reviews close, without any

communication) and sends his submission to the next journal on the list. Editors are used to this sort of behavior, and the less direct communication John has with editors, the safer he feels.

*Some people have speculated that paper mill submissions might benefit in peer review through bribery of journal editors. This is **extremely** unlikely, because the system is not overtly corrupt. Propose a brown paper envelope to an editor with a modicum of integrity, even at a journal with non-existent quality control, and you are toast. If there were academics out there normatively reaching out to editors with bribes, we would have heard more about it¹².*

John has another look at his ongoing commissions, trying to get an overview of which journals he is currently engaged with. He quickly finds one of his favored journals that he does not have a submission with at current.

In John's lab he quickly learned about journals with a reputation for low standards. During lunch breaks the postdocs would crack jokes about a perceived competitor lab and how their latest publication had been published in a 'crappy journal'. Can you believe professor this-and-that would stoop so low? Cue laughter. Even before John had been a part of his first publication, he would join in on the jokes, only half grasping the topics discussed. Little did he know back then that he was absorbing important information.

Learning the ins and outs of the landscape of scientific journals that publish the type of science you are engaged in is a part of every grad student's coming-of-age journey. But, unlike the other grad student in John's lab, he knows the landscape much better, almost like the back of his hand. Whereas he has less than five publications to his name, the true number of publications he has submitted is bigger. Much bigger.

Whenever something from John's lab is to be published in a journal, a meeting is held and a publication strategy is discussed. This is the one thing even the busy senior academics, like John's supervisor, will make time for. The contents of the upcoming publication is gauged for impact and topic. Was the work basic or applied science? Different journals focus on one or the other. What is the impact? That is, the perceived impressiveness of the experiments – how much was on display? Had something similar been done before? How well did the results conform to a hypothesis?

The right focus and impact means a highest tier journal – considered 'highest' tier mostly based on their impact factor, but also skewing reverently towards old and well-establishing journals in the field - followed by two or three tiers in descending order. If you have negative/null results? Back to the lab! But, in all other circumstances, the seniors at John's department will recommend to "aim for the stars", even when there is a low chance of success. One thing that scares a certain

¹² Notable exceptions are discussed in the 2013 Science exposé "[China's publication bazaar](#)" by Mara Hvistendahl.

breed of senior scientist is publishing in too low ranking journals. Having a paper accepted in the first-choice journal would make John's supervisor uneasy. Had he underestimated the contents? Could it have been published somewhere 'better'?

Aiming too high would, for a grad student or post doc, mean facing multiple rounds of rejection and hours spent reformatting references, fixing figures, writing cover letters and other time sinks related to journal submissions. But it would be no skin off the seniors' backs. They could only lose prestige by aiming too low. Shooting for the stars is easy as long as you are part of the ground crew. Things are a bit different onboard the rocket.

John opts for the exact opposite strategy. Using his knowledge of the journals in the field, and their reputation, he has dual aims: (1) publish somewhere 'good enough', with (2) as little friction as possible.

John is disappointed that one of his go-to journals would give him a picky reviewer with actual opinions on his text. Maybe it was a miscalculation on his part? Perhaps the journal has suddenly grown aspirations? Maybe they have a new editor? Maybe just bad luck. He makes a mental note to treat this journal more carefully in future. Strike one. If this becomes a pattern, the journal can't expect more submissions from John. He moves on with the second alternative.

The journal John decides to go for also comes from his mental shortlist of journals who have trouble filling out their (virtual) issues. Journals suffering from low submission volume sometimes accept most everything that comes their way. Journals that his colleagues would make jokes about. Journals that his research group would only bury unsuccessful manuscripts in, after a worn-down grad student had first desperately tried to submit it to seven other "higher tier" journals. John has submitted his work to them before, with pleasing results. He has the reference style ready to go in Zotero, and with a couple of clicks his references are good to go. The figures are ready as-is. John always exports his graphs and figures in a couple of formats that he knows covers his go-to journals. He has a template for the cover letter where he only needs to replace the name of the paper and the date of submission. After some shuffling in folders and some cursing at a journal submission system, his paper is off to the next journal on his mental list. In a little over an hour, he can go back and edit the status of the commission in his Excel sheet.

John will not notify the client about the change of plans unless they contact him. He has informed them - clearly, in multiple messages - that this part will take some time. He has also not pledged to submit to a particular journal, but rather to a tier of journal.

John's clients also work with an idea of tiers of journals. But unlike John's supervisor, his clients' tiers are less arbitrary. Rather, they are government or university mandated. And they do not mean ephemeral success or bragging rights, they are deadly serious. Here we need to understand a little of the motivation of John's clients.

Most paper mill papers that have been caught are ostensibly authored by researchers in South and East Asia. This seems to be tied to an overall move in these countries to "elevate" the level of research. This has been taken to mean that the model for research should be patterned on the process in North American and European universities. The metrics have consequently been based on publishing in high impact, English language journals ... with very little guidance on how this is supposed to happen.

To "incentivize" academics, everything from workplace promotions to straight-up cash bonuses have been tied to prestige publishing. Lecturers, teaching hundreds of students on a daily basis and with no time to carry out research, have been locked out of promotions, with no teaching-track modes of academic advancement offered. Clinicians, with no interest in research whatsoever, have been locked out of promotions unless they can demonstrate a glamorous publication list.

This is, of course, astonishingly unfair on the researchers themselves. They are the victims of systems of promotion and advancement which demand an effort that cannot realistically be achieved. And, of course, when other people meet that need by buying John's papers ... what are they to do? Science is glacial. John is fast.

Many of the people affected by these changes already have experience with paid English language services - proof reading and similar - to publish accurate, well-conducted, groundbreaking studies. It has become a trope for poorly written papers (or even well-written papers with non-Anglo names) to receive the review comment 'consult a native English speaker'. Scientific publications have become a nexus whereby capital is exchanged by multiple parties within multiple businesses. John's business is not so different in this aspect.

There are organizations listing the perceived prestige of English-language journals according to complicated, and opaque, formulas. These have come to play a crucial role in the focus on prestige publication. A one-time conversation over coffee with a visiting researcher from a South Asian university proved enlightening. Peer-reviewed journals are listed, by their impact factor, within their research field, by Scopus (a citation tracking database owned by Elsevier) into quartiles. Funding for the project that motivated his visit was granted, by a government scheme, contingent on the results being published in a journal that had a Scopus Q2 rating or higher. The project, the relieved researcher explained, was a success. But it does beg some questions. What do you do, in this situation, if your research is within a narrow field, and the two Scopus-Q2-or-higher journals reject your paper because your findings were negative? And you already spent the up-front funding?

Misaligned incentives and academics/clinicians put in impossible situations feed John's business. Whereas most steer clear of John and his ilk's services, no matter the situation, there are academics of every nationality who see research purely as a vehicle for their own promotion.

It would be so much easier if they'd settle for predatory publishers, though, John mutters to himself. He could churn out publications at ten times the speed. Having to deal with the vagaries of peer-reviewed - actually peer-reviewed - publishing is exhausting. Even journals desperate for submissions - barely in Scopus Q2 - cannot guarantee completely uncritical reviewers. They probably try very hard, but reviewers are in short supply these days. And if you shotgun out review requests, you get a lot of duds, but you also hit upon the occasional well-informed and critical voice.

Most of John's colleagues, who spend their lunch breaks joking about low quality journals, would consider themselves above accepting reviewing duties in the journals in which John publishes his faked science.

Here is the big-picture explanation to why John cites his own papers. Tanking the impact factor of his go-to journals by filling them with papers that are never cited (they are crap – make no mistake – so any serious author is unlikely to cite them other than in a hurry) bites him in the ass in the long run. Having them plummet in the rankings makes them unattractive to his clients, and means he has to find new outlets. And he would rather not. Moreover, an editor that might start to recognize a pattern, he figures, is more likely to turn a blind eye to his production as long as it serves to maintain (or even elevate) the journal's impact factor¹³.

Before he wraps it up for the day, John checks the status of four more papers. One of them is in the last stages, and John quickly approves the galley proofs of the publication. Many academics would celebrate this moment. For John, it mostly means he can cash in the other half of his client's payment. He will celebrate the PayPal transfer instead.

We don't know how money changes hands when it comes to John's business, but it can be assumed that no one would be willing to pay him the full amount up front. Grant writers don't work like that, neither do authors, neither does anyone working on a research contract, or any other similar agreement. Remember, John cannot advertise in the New York Times. He wants to be the most prolific paper mill you have never heard of. We can assume he's referred primarily through word of mouth. With only one or two people recommending a service, would you be willing to pay cash up front? Similarly, it seems unlikely that John would do the heavy work of faking together a complete report and submitting it for a client who has transferred no funds, and might still back out of the deal. Of course, he can sell the paper elsewhere, but then his clients have no buy-in, make no commitment. Consequently, it seems likely that money changes hands at two or more points in the process. Or, perhaps Bitcoin¹⁴.

¹³ A journal's impact factor is calculated as the number of times its papers are cited, on average, during the first two years following their publication.

¹⁴ Industrious people offering to write you an essay have been known to use the Chinese messaging app WeChat which has integrated money transfer options. Outside of China, financial transactions have periodically been blocked, why this is probably not an app our paranoid John would rely on.

In some cases John might be a bit more trusting. Sometimes his clients might not be on the author line of his papers at all. Manuscript services catering to non-English-speaking authors wanting to publish in English-language journals (as demanded by the government institutions they work at) have been known to branch out. Starting with offering language checks and translation services, some have branched into increasingly dubious services. Some are even offering publications with 'without doing experiments'/'no need to do any research'. Whereas these organizations employ writers fluent in English, they cannot provide the services John deals in. So, in order to accommodate their clients with no time to do actual research, they hire John. In these cases, where his client is in as deep as is John, and they have a symbiotic relationship, John might be less paranoid about letting his client know his identity and work with one-time payments.

As John clocks out, turning off his computer, he thinks a bit about his future. He remembers when his bank statements were gruesome. Then, suddenly, they were stable, and then after that – against all odds, for a young scientist – strongly in the black. Perhaps he could drop out of his current position at the university and do this full time? He has considered the economics of his situation before. It is feasible. His biggest problem would be social: how to explain to friends, family, and former colleagues what he does for a living? Scientific consultant? Medical writer? The truth, at absolute minimum, would make them extremely uncomfortable.

But there's always a way to avoid awkward questions. Obfuscate, change a few details, use a few synonyms, and hide in plain sight. He is good at that.

Maybe it would work.

Maybe it could stay a secret.

PART TWO: BREAKING THE MILLSTONE

By James Heathers

How could John exist?

The short answer is: because every system is exploitable, and science relies on a form of *social* rather than *procedural* policing. And because advantages that exist within constrained systems are eventually taken.

The long answer is a lot longer.

Science has a weird undercarriage built on trust. While people like me often complain bitterly about the hypocrisies which arise around that trust, it is a real thing, and it does pose and reinforce mutual values. This gives rise to *social policing*, where mutual communications and expectations about the structure of the scientific enterprise maintain its own internal structure.

Imagine Fraud A meets Fraud B. They know each other's work well, and have strong suspicions that they share the same ... 'flexible' approach to engineering scientific publications dishonestly. But the environment is such that *they would both feel very uncomfortable discussing it with each other*. In game theoretical terms, this is called *mutual knowledge*. The situation is reminiscent of the atheist priests described by Dennett and LaScola¹⁵:

One can be initiated into a conspiracy without a single word exchanged or secret handshake; all it takes is the dawning realization, beginning in seminary, that you and the others are privy to a secret, and that they know that you know, and you know that they know that you know... Without any explicit agreement, mutual knowledge seals the deal: you then have no right to betray this bond by unilaterally divulging it, or even discussing it.

Incidentally, this is why I distrust surveys about scientific misconduct. Even if they are completely anonymous, you are expecting a dishonest person to honestly report their dishonesty, and in doing so, throw a marble into a bucket that will attract a colossal amount of unwanted attention if it fills. Frauds may not understand the computational back-end of what they create, they may leave very obvious clues in their data (if you know how to find them), but there is no reason for them to be *socially* stupid. They have, for instance, understood how not to get caught so far. Even if they can't add.

Social policing works when everyone shares the same values. When they don't, we fall back on explicit rules and real policing – the enforcement of regulations by statutory bodies designed

¹⁵ https://ase.tufts.edu/cogstud/dennett/papers/Preachers_who_are_not_believers.pdf

to accept reports of, investigate, and prosecute violations. In the absence of that, and make no mistake *we do not have it, no matter what anyone says on the internet*, a system is entirely available to be manipulated. If science had a direct business analogue, everyone from two-bit crooks with gold teeth to hedge funds would be strip-mining it blind. There is an upside: some local fame and moderate fortune. There is a catastrophic downside, but one which is easily hedged against with straightforward risk mitigation.

Anyway. If Otto gets to tell a story, then so do I. Here's my direct analogy to scientific publishing:

The government is stockpiling an astonishingly large amount of microchips, because they are 'useful'. So, hordes of manufacturers just turn up at the gates of infinite warehouses, and sell lorryloads of them. The sale is handled by the guard on the gate.

*The quality of the guard is generally reflective of the quality of the warehouse – which ranges from good, to OK, to absolutely shithouse. Sometimes the guard is fully awake, and inspects the contents of the truck. He opens the boxes, opens the packaging, pulls out the chips, pops them in a test circuit. Defective devices are not purchased. Elsewhere, the guard is fairly lax – opens the lorry, sees plain brown boxes with **TOTALLY CHIPS INSIDE I PROMISE** written on them in Sharpie, and thinks 'good enough'. Elsewhere still, the guard is almost asleep, and will just wave the truck in for unloading, throwing a receipt for goods through the delivery driver's window without looking up from a copy of Newsweek.*

The better warehouses pay more, but you can make a living from selling to the bad ones.

*Regardless, one rule applies to all of them: if you try to cheat the guard, sneak in the wrong product, the guard won't take your product, but they will invariably let you go with either zero consequences or a stern warning. Only on the rarest occasions will this polo-shirted hero pick up the phone and call the number on the side of the truck: "Did you know your rep just tried to sell me bricks with **MIKROCHIPPE** written on the side?? Fire this clown! No, not from his job, right into the sun!"*

Not only that, but if you have a lorry full of empty boxes or birdseed or stale cookies, and the guard asks if you want to hang around for an inspection, you can just say no, that's fine, I'll leave – and he won't pursue the issue further. Not his problem.

The whole system functions because the good microchip manufacturers care desperately about both their reputation and the mission of the government as a whole. They want to build a good product, are generally happy to have it inspected, are quite interested if the inspection points out problems they can improve, and quite satisfied when their product is used to build things elsewhere.

John isn't like them. He is a clever middleman, who works for the distributors. He has a business selling empty cardboard boxes with "T089-83 SuperCore" written on the side, along with some circuit diagrams and a cool graphic of lightning coming out of a PCB. Then he weights them with gravel and packing foam, until they feel nice and solid. This is not instant, but it certainly is

lightning quick compared to actual manufacturing. Unopened, the boxes look and feel a lot like they're full of microchips, much better than the people who write COMPUTA PARTS on the side of the box in crayon.

Through careful titration, John has a shortlist of favoured warehouses, at the intersection between the maximum quality of warehouse that will accept the goods with the minimum amount of inspection. He does not make money when his shipments get rejected, because he works for the distributor to complete the sale – they hire John to fill their trucks with boxes, drive them to the warehouse, get the shipments accepted, and get them paid. Only then does John get paid himself.

From this stockpile, the government uses some of the microchips, some of the time. When they need a new batch, they go to the top warehouses, where there are site inspections and quality control, and draw some of their chips from the stockpile. They work, mostly.

But John's warehouses full of empty boxes just sit there, uninspected, maybe forever, while the distributors get paid and John gets his cut. The risk is low. It's actually quite hard to get the guards to do an inspection at all, and when they do, they often open the warehouse door, look at the boxes, and mutter 'yeah that's fine'. And go straight back to their Newsweek, which has a particularly engaging story about scientific fraud this week.

I cannot be more explicit about this: make a place where people can grift, and grifters will appear. It is the only reliable place in the human experience where the phrase 'build it and they will come' holds as accurate.

There is a great story in an old David Kushner book: a small and somewhat obscure casino once had a celebration, and offered 2/1 odds on blackjack instead of 3/2 – a fun little bonus for their regular clientele. But when the tables open for the night, the character of the book – a professional cardsharp, an odds-beater – looks around the casino, and realizes: *everyone else* is also a shark. The whole casino, every table, was filled with hard players who could count cards as easily as breathing. The odds were already slightly in their favour, but this new rule, perhaps a small detail to us, meant that today the casino had made themselves the sucker. Within the space of a few hours, the casino got mauled. "Build it and they will come" only works when you offer your neck - a certain sort of person will turn up to bite it.

Now, science is not a business in any straightforward sense. Often, it has lore rather than explicit rules, and we rely on ostensive gatekeeping values to ensure that people who do not have human progress and curiosity as core motivations are squeezed out. Hell, sometimes that feels like the only reason to stay. As a human enterprise, it has evolved an odd combination of circumstances around publication (enforcement, quality control, reward, competition/pressure, and available resources) that have produced a system not designed to detect this class of exploit.

So, John is the exploit, and he's provided a service that's priced low enough, in sufficient demand, and delivers the desired outcome.

We shouldn't be too surprised. In some ways, he's a step in a continuum of services that are provided in scientific production. Just within domains that I've heard of, you can hire a grant writer for grant text, a grant manager or consultant for grant back-end/budgetary/government requirements, a ghostwriter for papers, a statistician who will be less than scrupulous with your data for analysis, a computational consulting service that will comb through *every* possible analytical pathway in your data looking for reportable results (*i.e.* a multiverse p-hacker), an 'editing service' for manuscripts, and those are just the experts that people might admit to hiring.

While these services differ in their cost, and perceived/actual dishonesty, the one thing they all have in common: they aren't free. For most researchers, hiring a grant writer is an impossible dream, or likewise working at an institution which has one available. It can be an astonishingly good investment: think about what your return is on a \$50k grant writer salary if you apply for multiple \$1M-\$5M grants in a year. Discretionary or departmental budgets most likely pay for these, as a grant agency will not appreciate you asking for their money in order to pay someone to ask for their money. Companies in research agreements and/or partnerships may as well.

To make that work, of course, you must have access, money and connections. Just like everything else, everywhere else, we see preferential attachment at work¹⁶.

This means John's service is actually perversely democratic – it's accessible to anyone with some personal funds if they know it exists. I'm not cynical enough to presume that departmental budgets are spent on John. You might admire the accessibility, if it wasn't completely dishonest, desperately cynical, and potentially illegal.

So, that's how John can exist.

But, does he?

We think so. But how to prove it? It is one thing to have suspicions, it is quite another to have evidence. This is a narrative constructed from hints, suggestions, frustrations, and suspicions.

We don't know much about them. Where they are, whether or not they're a company, how many employees there are if they are, how many people there are doing this, how many of these are getting bought, where they're all going, what subfields they end up in...

¹⁶ This is how we say 'the rich get richer' these days. Language evolves, even if people don't.

That's just us, though. There's *some* evidence already. See Footnote 3 in Otto's piece above, if you haven't already.

Those investigations are all worth your time, because they posit a network or an ecosystem of John-like actors, and they present good evidence. Our evidence is more specific to a single case, but also maddeningly abstract. There isn't enough proof to outline specific charges, which could be premature and unfair to authors, and also... frankly we don't want to tip our hand. John can adjust his service quality upwards, make the template just a little bit better, and scam us all forever *if* he knew what to change.

And he might read this.

Hi, John. You irredeemable sack of shit.

How ironic. Writing about a system being exploited for its reliance on trust, and making a claim we're unable to fully outline, relying itself on trust.

Where does he work?

Largely: small journals, fringe journals, pay-to-play-and-probably-used-to-be-on-Beall's-List but still somehow good enough to be indexed journals. And, of course, the 'used to be a real journal but fell on hard times and/or incredibly poor stewardship' journals. Make no mistake, outrageous cookie-cutter fraud is probably *only slightly* more common than you think it is within the pages of the brand-name journals.

But in the enormous undifferentiated rump of journals you've never heard of, this is not the case at all. These are the paper repositories which give rise to the figures like "a third of academic articles are never cited", or whatever it is now. These are largely unread documents which form a line in a CV, exist only to benefit the author, and are too obscure, poorly indexed, or instantly untrustworthy to amount to consequential information for other people.

Make no mistake: when you see a journal with an impact factor of 0.5, you can reliably assume that the bulk of articles are never cited. Articles in eLife a few years ago had a *median* citation rate of 6, but at +2SDs, it was more like 28 citations, with a maximum of 136¹⁷. Think of that same left-censored right-skewed curve from the linked article, and apply it to a journal with a low profile by squeezing the whole curve leftwards.

Basically, no-one is paying attention to those. It is a scandal when a Nature paper is retracted. It's in the newspapers. Less people are aware of, say, the fact that The Journal of Biological

¹⁷ <https://scholarlykitchen.sspnet.org/2016/07/18/will-citation-distributions-reduce-impact-factor-abuse/>

Chemistry has more than 400 entries in the Retraction Watch database¹⁸. That's not just a few rotten apples. The whole orchard has Irritable Bowel Syndrome, and has shit itself with power and authority.

Summing up

There are, of course, consequences to this collective devaluation of the scientific literature, and I see all this getting worse before it gets better. Right now, people with careers, promotions, jobs, etc. resting largely on papers that they bought are out there giving talks on that same work.

Something has to change. If it's possible to be sunny for a moment - it might be easier than you think. Enough pressure on John to produce better differentiated papers, more sophisticated frauds, and the cost of effort will increase. At some point in time, the pool of people with sufficient private funds to engage a more sophisticated service might get squeezed.

Maybe.

But it's just as likely this will result in not market pressure, but equally sophisticated changes to the process. More authors per paper, paying into the same service. NLP-based paper generators. Artificially generated unique images – no more image duplications to spot. Reverse-engineering the techniques used to spot garbage, and avoiding them. If we have more sophisticated cats, we may just get a mouse who can produce paper narratives using Pytorch.

What else would help, of course, are the same things that everyone in the Open Science community have been banging on about for years. Open review – find out who's admitting these papers into record. Data – send it. Data is harder to fake than results, and if you do it the same way twice, we can find a pattern. Methods, code – include them in full. At some point, they'll contradict the results, or themselves. Jupyter or online notebook style papers have higher demands for skullduggery – so use them. Post publication review – take it seriously. Invalidations of previously published work – actually act on them.

There's no point, though, in even expanding the above into full sentences. You've heard it all before, as have I, and you'll go blue holding your breath waiting for these to be widely implemented. Support in favour of these ideas is trudging along, of course, and that is excellent. But in academia, a good idea can wait outside the castle gates and starve for decades.

Now, what do I think will really move the needle?

The honest answer is: a massive scandal. A complete blood-letting. A medieval living autopsy of something so brazen and bastardic that it makes the news, even while the infectious, burning,

¹⁸<http://retractiondatabase.org/RetractionSearch.aspx#?jou%3dthe%2bjournal%2bof%2bbiological%2bchemistry>

screaming world consumes itself. The full delineation of an act so ridiculous that it actually manages to knock some rust scale off the apathy.

So, we'll see. To be continued.

You're out there, John. But we're out here, too. And, when I find your neck, I will squeeze it.