AUTHENTICATED
U.S. GOVERNMENT
INFORMATION
GPO

PAPER MILLS AND RESEARCH MISCONDUCT: FACING THE CHALLENGES OF SCIENTIFIC PUBLISHING

HEARING

BEFORE THE

SUBCOMMITTEE ON INVESTIGATIONS AND OVERSIGHT

OF THE

COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

OF THE

HOUSE OF REPRESENTATIVES

ONE HUNDRED SEVENTEENTH CONGRESS

SECOND SESSION

JULY 20, 2022

Serial No. 117–65

Printed for the use of the Committee on Science, Space, and Technology

Available via the World Wide Web: http://science.house.gov

U.S. GOVERNMENT PUBLISHING OFFICE

48–028PDF WASHINGTON : 2022

COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY

HON. EDDIE BERNICE JOHNSON, Texas, *Chairwoman*

ZOE LOFGREN, California
SUZANNE BONAMICI, Oregon
AMI BERA, California
HALEY STEVENS, Michigan,
Vice Chair
MIKIE SHERRILL, New Jersey
JAMAAL BOWMAN, New York
MELANIE A. STANSBURY, New Mexico
BRAD SHERMAN, California
ED PERLMUTTER, Colorado
JERRY MCNERNEY, California
PAUL TONKO, New York
BILL FOSTER, Illinois
DONALD NORCROSS, New Jersey
DON BEYER, Virginia
CHARLIE CRIST, Florida
SEAN CASTEN, Illinois
CONOR LAMB, Pennsylvania
DEBORAH ROSS, North Carolina
GWEN MOORE, Wisconsin
DAN KILDEE, Michigan
SUSAN WILD, Pennsylvania
LIZZIE FLETCHER, Texas

FRANK LUCAS, Oklahoma,
Ranking Member
MO BROOKS, Alabama
BILL POSEY, Florida
RANDY WEBER, Texas
BRIAN BABIN, Texas
ANTHONY GONZALEZ, Ohio
MICHAEL WALTZ, Florida
JAMES R. BAIRD, Indiana
DANIEL WEBSTER, Florida
MIKE GARCIA, California
STEPHANIE I. BICE, Oklahoma
YOUNG KIM, California
RANDY FEENSTRA, Iowa
JAKE LATURNER, Kansas
CARLOS A. GIMENEZ, Florida
JAY OBERNOLTE, California
PETER MEIJER, Michigan
JAKE ELLZEY, TEXAS
MIKE CAREY, OHIO

SUBCOMMITTEE ON INVESTIGATIONS AND OVERSIGHT

HON. BILL FOSTER, Illinois, *Chairman*

ED PERLMUTTER, Colorado
AMI BERA, California
GWEN MOORE, Wisconsin
SEAN CASTEN, Illinois

JAY OBERNOLTE, California,
Ranking Member
STEPHANIE I. BICE, Oklahoma
MIKE CAREY, OHIO

C O N T E N T S

July 20, 2022

Page

Hearing Charter 2

Opening Statements

Statement by Representative Bill Foster, Chairman, Subcommittee on Investigations and Oversight, Committee on Science, Space, and Technology, U.S. House of Representatives 9
Written Statement 10

Statement by Representative Jay Obernolte, Ranking Member, Subcommittee on Investigations and Oversight, Committee on Science, Space, and Technology, U.S. House of Representatives 12
Written Statement 12

Written statement by Representative Eddie Bernice Johnson, Chairwoman, Committee on Science, Space, and Technology, U.S. House of Representatives 13

Witnesses:

Dr. Jennifer Byrne, Director of Biobanking, New South Wales Health Pathology; Professor of Molecular Oncology, University of Sydney
Oral Statement 15
Written Statement 18

Mr. Chris Graf, Research Integrity Director, Springer Nature; Chair of the Governance Board, STM Association Integrity Hub
Oral Statement 25
Written Statement 27

Dr. Brandon Stell, Neuroscientist, French National Centre for Scientific Research; President and Co-Founder, The PubPeer Foundation
Oral Statement 36
Written Statement 38

Discussion 43

Appendix I: Answers to Post-Hearing Questions

Dr. Jennifer Byrne, Director of Biobanking, New South Wales Health Pathology; Professor of Molecular Oncology, University of Sydney 56

Mr. Chris Graf, Research Integrity Director, Springer Nature; Chair of the Governance Board, STM Association Integrity Hub 59

Dr. Brandon Stell, Neuroscientist, French National Centre for Scientific Research; President and Co-Founder, The PubPeer Foundation 63

Appendix II: Additional Material for the Record

Documents submitted by Representative Bill Foster, Chairman, Subcommittee on Investigations and Oversight, Committee on Science, Space, and Technology, U.S. House of Representatives
"Neutron Production and Absorption in Uranium," H.L. Anderson, E. Fermi, and Leo Szilard 66
"Uranium: Neutron Production and Absorption," B. Foster and E. Perlmutter 71
Letter, Allison C. Lerner, Inspector General, National Science Foundation 74

PAPER MILLS AND RESEARCH MISCONDUCT: FACING THE CHALLENGES OF SCIENTIFIC PUBLISHING

WEDNESDAY, JULY 20, 2022

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON INVESTIGATIONS AND OVERSIGHT,
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY,
Washington, D.C.

The Subcommittee met, pursuant to notice, at 10:02 a.m., in room 2318 of the Rayburn House Office Building, Hon. Bill Foster [Chairman of the Subcommittee] presiding.

U.S. HOUSE OF REPRESENTATIVES
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY
SUBCOMMITTEE ON INVESTIGATIONS AND OVERSIGHT

HEARING CHARTER

Paper Mills and Research Misconduct: Facing the Challenges of Scientific Publishing

Wednesday, July 20, 2022
10:00 a.m. EDT – 12:00 p.m. EDT
Zoom

PURPOSE

The purpose of this hearing is to discuss the current and future challenges in securing scientific literature from fraudulent academic papers. The hearing will examine field-specific and industry-wide strategies for identifying fraud, the increasing number of fraudulent papers produced and sold by paper mills, and the impact of new technologies such as AI on both the perpetration and the detection of research misconduct. Members and witnesses will discuss the successes of the largely volunteer post-publication review community, the challenges that community has faced, and the strategies publishers themselves are developing to combat research misconduct.

WITNESSES

- **Dr. Jennifer Byrne**, Director, Biobanking, New South Wales Health Pathology; Professor of Molecular Oncology, University of Sydney
- **Mr. Chris Graf,** Research Integrity Director, Springer Nature; Chair, Governance Board, STM Association Integrity Hub
- **Dr. Brandon Stell,** Neuroscientist, French National Centre for Scientific Research; President and Co-Founder, The PubPeer Foundation

OVERARCHING QUESTIONS

- What is the scope and what are the sources of misconduct in academic publishing?
- What tools and methodologies exist to detect research misconduct both prior to and following publication?
- How are these tools being deployed and who is responsible for deploying them?
- How will automation help or hinder the fight against research misconduct in the scientific literature?

BACKGROUND

What is Research Misconduct?

In December of 2000, the Office of Science and Technology Policy (OSTP) defined research misconduct as "fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results," harmonizing conflicting definitions across the US Federal scientific enterprise.[1] Fabrication refers to creating data out of whole cloth, falsification is manipulating results to give a desired outcome, and plagiarism is the appropriation of another's work. Any of these alone is sufficient to identify an action as research misconduct, and all fall under the broad umbrella of academic fraud.

The OSTP definition, and the subsequent investigative procedures, specifically covered research conducted with Federal funds. However, since research misconduct renders the conclusions of research invalid, no papers which are the product of misconduct should be present in the scientific literature, whether or not they were supported with public funds. Ideally such papers would be identified early and rejected from the publishing process, either during peer review or during evaluation by research integrity professionals employed at the journals themselves. Unfortunately, researchers studying academic fraud identify numerous published papers which contain evidence of research misconduct every year.

Why commit research misconduct?

Individual research misconduct is when a researcher or researchers commit falsification, fabrication, or plagiarism to make their work publishable. Techniques for committing this fraud can vary by field. For instance, in the life sciences key data often takes the form of images such as cellular assays or western blots, which are relatively easy to digitally alter to better reflect a desired result. Other fraud might include altering datasets by adding or removing data to better make it demonstrate a desired or significant result.

In academia there is a concept colloquially referred to as "publish or perish," i.e., that a certain quantity of publications is required to advance in a career. In the U.S., the number of publications, the number of citations on those publications, and the impact factor of the journal that prints a researcher's work are common metrics to assess a researcher's quality as a scientist and their worthiness for recruitment or prestigious teaching positions.[2] Conducting research is just one among many responsibilities of most academics, and even when performed well, publication-worthy results are not guaranteed. As such, there is an incentive to supplement output by using research misconduct as a shortcut. Recognizing this problem, some organizations have taken steps to standardize alternative metrics for assessing research quality.[3] For instance, the San Francisco Declaration on Research Assessment pushed for a greater focus on the quality of research rather than the number or location of publications.[4]

---

[1] https://www.govinfo.gov/content/pkg/FR-2000-12-06/pdf/00-30852.pdf
[2] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3999612/
[3] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2126010/?tool=pubmed
[4] https://sfdora.org/read/

Incentives for research misconduct can be even stronger in other countries. In Russia, the 2012 May decrees included a push for certain universities to increase their international ranking, in part by increasing their number of publications.[5] In China, physicians were required to publish research articles, often in English language journals, to gain promotions or even be permitted to perform certain surgeries, regardless of whether they possessed the time or resources to conduct such research.[6] China has taken steps to remove explicit cash incentives for publishing low quality papers,[7] but there are still strong incentives to publish a paper, any paper, to move to the next level in a career.[8] So long as the academic community, both domestic and international, values publication quantity over quality, the incentive to commit research misconduct will remain.

Paper Mills

The challenge of scaling research misconduct oversight has grown along with the incidence of "paper mills." Paper mills are organizations, often based in Russia or China, which sell scientific papers, authorship positions on papers, or other fraudulent academic products.[9] They may disguise themselves as offering translation services for non-native English-speaking researchers.[10] Some will create a fake email account and entirely take over the correspondence with journals on behalf of the customer. One research integrity expert told the Committee that it is a strong indicator of a paper mill when a non-native English-speaking researcher is too responsive or too willing to share data or other information to move along their paper submission. While some use sloppy and unprofessional formatting, others are careful to match the submission standards of the journals to which they send their fraudulent papers.

Paper mills have a multitude of strategies to source the fraudulent papers they sell and/or attempt to publish for a price. One technique is to contact researchers with offers to pay to add fake co-authors to legitimate research. Another strategy is to take over the editing process by bribing editors to turn a blind eye to fraudulent papers in the submission queue or by purchasing editorial control from unscrupulous editors.[11] They have also been known to hijack entire journals either by assuming the digital identity of the journal,[12] or by using identity fraud to pose as a guest editor of a journal to guarantee that their own papers are accepted.[13] Paper mills may also create their own papers en masse through ghostwriting, templates, or the application of automated tools that will be discussed in more detail below.

No complete census of these organizations exists, and the scope of their impact is difficult to determine. But anecdotally, independent researchers who investigated one Russian paper mill were able to identify hundreds of papers it had produced and published, leading to the retraction

[5] https://arxiv.org/ftp/arxiv/papers/2112/2112.13322.pdf
[6] https://www.nature.com/articles/d41586-021-00733-5
[7] https://www.nature.com/articles/d41586-020-00574-8
[8] https://publicationethics.org/resources/forum-discussions/publishing-manipulation-paper-mills
[9] https://www.science.org/content/article/russian-website-peddles-authorships-linked-reputable-journals
[10] https://publicationethics.org/resources/forum-discussions/publishing-manipulation-paper-mills
[11] https://retractionwatch.com/2018/05/30/want-to-earn-10k-per-month-join-the-journals-mafia/
[12] https://retractionwatch.com/2022/05/29/want-to-know-whether-that-journal-is-scamming-you-introducing-the-retraction-watch-hijacked-journal-checker/
[13] https://retractionwatch.com/2021/06/18/galling-journal-scammed-by-guest-editor-impersonator/

of 30 papers from one journal alone.[14] Another paper mill based in China, dubbed "the tadpole paper mill" for the shape of the fraudulent western blot images it produced, has been linked to more than 400 papers published primarily across six journals.[15]

ENTITIES COMBATTING MISCONDUCT IN ACADEMIC PUBLISHING

Publishers

Academic publishers have a vested interest in protecting the reputations of their journals by detecting fraudulent articles prior to publication. Peer review, in which subject matter experts are recruited to examine a paper within their discipline for worthiness, is the classic quality control method of academic publishing. However, it is unreasonable to expect subject matter experts to also excel at forensic examinations of images, data, and author backgrounds. As such, checks for research misconduct must occur during other phases of the publishing process, an effort which requires its own dedicated resources.

Publishers have several avenues for cooperation to help detect and mitigate misconduct in the papers they publish. One example is the Committee on Publication Ethics (COPE), which was formed in 1997 to coordinate on misconduct cases.[16] COPE issues guidance to publishers, provides example protocols for handling research misconduct, and creates flowcharts for specific fraud types such as image manipulation[17] or systemic fraud.[18] COPE began discussing paper mills in earnest in the fall of 2020[19] and, in a partnership with the International Association of Scientific, Technical and Medical Publishers (STM) released a paper analyzing paper mills in depth in June of 2022.[20]

STM has also begun developing a collaboration platform for the 140 publishers that comprise their membership. This platform is called the STM Integrity Hub and Chris Graf chairs the governance board. The first major component of this project will be to develop a system that will detect simultaneous submissions to multiple journals – a strong indicator of paper mill activity – while protecting the intellectual property of the publishers and the privacy of paper authors.[21] Future efforts will aim to facilitate the development and exchange of research integrity best practices and the creation of additional tools to detect research misconduct.

Volunteer Post-Publication review community

Despite the efforts of publishers, many fraudulent papers are detected following publication by a primarily volunteer post-publication review community. Many members of this community started as academics and then altered or expanded their trajectory after a serendipitous encounter with research misconduct. For instance, Dr. Elisabeth Bik was an academic microbiologist by

[14] https://retractionwatch.com/2022/07/05/our-deepest-apology-journal-retracts-30-likely-paper-mill-articles-after-investigation-published-by-retraction-watch/
[15] https://scienceintegritydigest.com/2020/02/21/the-tadpole-paper-mill/
[16] https://publicationethics.org/about/history
[17] https://publicationethics.org/files/image-manipulation-published-article-cope-flowchart.pdf
[18] https://publicationethics.org/files/publication-process-manipulation-cope-flowchart.pdf
[19] https://publicationethics.org/resources/forum-discussions/publishing-manipulation-paper-mills
[20] https://publicationethics.org/files/paper-mills-cope-stm-research-report.pdf
[21] https://www.stm-assoc.org/stm-integrity-hub/

training who now makes a living through her expertise at detecting image manipulation and her services as a research integrity consultant.[22] Others, such as Dr. Jennifer Byrne, are academics who have broadened their area of study while maintaining a traditional professorship at a research university. While she eventually received a small grant from the Office of Research integrity, Dr. Byrne's work scanning papers for wrongly identified nucleotide sequencing reagents was initially undertaken without funding.[23] Still other volunteers examine raw numerical data for signs of manipulation using statistical tests.[24]

These volunteers, many of them anonymous, conduct ad-hoc reviews based on their respective interests and skills. There is no overarching organization independent of journals themselves to fund fraud detection or target it to where it is most needed within the scientific literature. The quantity of papers detected by volunteers lacking a systemic approach suggests they are in a target rich environment, i.e., the full scope of fraud within the scientific literature is larger than is currently known. Unfortunately, some volunteers who identify research misconduct report harassment and lawsuits from authors they have called out,[25] and others stay anonymous in part to avoid receiving that kind of reaction.[26] Other volunteers reported giving up on notifying journals of their findings after repeatedly being stonewalled while clearly flawed papers were allowed to stand unchallenged.

To combat fraud as effectively as possible, these volunteers have coalesced around several websites for communication and coordination, including Twitter.[27] The blog Retraction Watch is also a valuable resource for the community as a provider of the latest news on fraud detection, a place to post findings that might not fit in traditional journals, and as a database of all journal retractions.[28] Another website, PubPeer.com – founded by Dr. Brandon Stell – provides a place for anyone to comment on concerns on scientific articles published in any journal. The website provides a space for authors to discuss potential problems in their work directly with the people who detected those problems.[29]

Federal Government

A finding of research misconduct could end a scientist's career, so Federal science agencies follow a structured process for the investigation and adjudication of allegations of scientific misconduct.[30] This process applies strictly to research funded by Federal agencies.

One Federal leader for detecting research misconduct is the Office of Research Integrity (ORI) housed within the Department of Health and Human Services. ORI's stated mission is to oversee all Public Health Service (PHS) research integrity issues.[31] It performs this mission in part by conducting investigations into cases of research misconduct using PHS funding and by providing

[22] https://scienceintegritydigest.com/about/
[23] https://www.nature.com/articles/d41586-021-02136-y
[24] https://www.science.org/content/article/meet-data-thugs-out-expose-shoddy-and-questionable-research
[25] https://www.the-scientist.com/news-opinion/elisabeth-bik-faces-legal-action-after-criticizing-studies-68831
[26] https://forbetterscience.com/2019/07/30/help-with-another-not-on-pubpeer-yet/
[27] *Ibid.*
[28] https://retractionwatch.com/the-center-for-scientific-integrity/
[29] https://pubpeer.com/static/about
[30] https://www.govinfo.gov/content/pkg/FR-2000-12-06/pdf/00-30852.pdf
[31] https://ori.hhs.gov/policies-ori-mission

small grants – between $50K and $150K – for research, development, and demonstration of research integrity projects.[32] ORI's investigations are thorough, in keeping with the OSTP guidelines from the definition of research misconduct, but they are also relatively few, averaging just 4 case summaries published each year since 2016.[33]

Though it does not offer research grants, the NSF Office of the Inspector General performs a similar function to ORI by conducting research misconduct investigations when there are credible allegations of misconduct in NSF-funded research. The OIG has tools to detect plagiarism. It also relies on tips from the public or reporting from the volunteer review community, including through Retraction Watch and PubPeer, to bring fabrication or falsification cases to their attention for evaluation.

Automation

Automated tools for translating or modifying text so that it can pass through a traditional plagiarism detector are being used to help to disguise plagiarism in scientific papers.[34] In 2005, graduate students at MIT developed a tool called SCIgen which could produce a nonsensical but well formatted computer science research paper on demand. Researchers using the tool managed to get several papers published in a variety of journals over the years as a pointed statement on journal vetting processes.[35] Another robo-writer tool is designed specifically to produce Small Business Innovative Research (SBIR) proposals. At this time, the results are generally jargon-heavy to disguise their fundamentally nonsensical nature but could easily be detected by an attentive human with relevant expertise. Recent developments have brought incredible strides in AI-powered language models. Tools such as GPT-3 show an incredible versatility with a range of writing formats, and robo-written scientific material will only grow more sophisticated along with these tools.[36]

However, while automation may be enabling more efficiencies for committing research misconduct, it also offers new opportunities for detection and mitigation. Automated tools can enable publishers and volunteer reviewers to evaluate papers more quickly and thoroughly than would be possible for a human working unaided. Note that even advocates stress that humans should always make the final determination of research misconduct when automated tools are used to assist with detection.

Detecting plagiarism: The detection of direct plagiarism has long been the focus of research misconduct tools. Millions of students have had their papers checked by Turnitin.com since the tool's creation in 1998, and publishers commonly employ a variant called iThenticate or other similar services to check for plagiarism in submitted articles.[37]

Researchers have discovered a way to detect the use of robo-writers through specific "tortured phrases" where pairs of phrases are technically synonyms, but a human would clearly see a

[32] https://ori.hhs.gov/blog/fy-2022-grant-opportunity-forecasts
[33] https://ori.hhs.gov/content/case_summary
[34] https://www.nature.com/articles/d41586-021-02134-0
[35] https://pdos.csail.mit.edu/archive/scigen/#people
[36] https://www.technologyreview.com/2020/07/20/1005454/openai-machine-learning-language-generator-gpt-3-nlp/
[37] https://www.turnitin.com/products/ithenticate

distinction. For instance, there is a clear difference between "breast cancer" and "bosom peril." Once a tortured phrase such as bosom peril is identified it can serve as an easily detectable marker of a plagiarized paper.[38] A study searching for tortured phrases found almost 8,000 in the published literature, many in reputable journals.[39]

Detecting fabrication: OpenAI, which creates the GPT series of AI-powered language models, has also produced a detector for identifying text created by the GPT-2 version of the tool. When researchers applied that tool to Elsevier paper abstracts, it flagged hundreds of papers for containing synthetic text likely produced by GPT-2.[40]

Detecting falsification: The tadpole paper mill was found in part because hundreds of fraudulent papers all used similar images with an eponymous tadpole-like structure[41]. Multiple companies, such as Image Twin[42] and Proofig,[43] are developing tools that will automatically detect manipulation or duplication of images across papers. Those tools have shown early promise, but they have not yet been externally evaluated or scaled up to the level needed to handle even the hundreds of fraudulent images known to exist.

Another tool called Seek & Blastn was created to detect highly field-specific fraud. It scans papers and extracts gene identifiers and nucleotide sequences from the text and flags potential incorrect usages of reagents and sequences for humans to review .[44] Using this tool, a team of researchers was able to reduce a population of 12,000 papers down to 3,400 with potential problems and ultimately identified 712 flawed papers.[45] This tool is limited in scope, as it cannot recognize non-human genomes, but it demonstrates that automated tools can provide valuable assistance even when the fraud in question requires significant scientific expertise to recognize.

[38] https://arxiv.org/abs/2107.06751
[39] https://dbrech.irit.fr/pls/apex/f?p=9999:24::IR_years
[40] https://thebulletin.org/2022/01/bosom-peril-is-not-breast-cancer-how-weird-computer-generated-phrases-help-researchers-find-scientific-publishing-fraud/?utm_source=Twitter&utm_medium=SocialMedia&utm_campaign=TwitterPost01132022&utm_content=DisruptiveTech_TorturedPhrasesPointToFraud_01132022
[41] https://scienceintegritydigest.com/2020/02/21/the-tadpole-paper-mill/
[42] https://imagetwin.ai/
[43] https://www.proofig.com/
[44] https://www.protocols.io/view/seek-amp-blastn-standard-operating-procedure-q26g7b2k1lwz/v1
[45] https://pubmed.ncbi.nlm.nih.gov/35022248/

Chairman FOSTER. Well, thank you. And this hearing will now come to order. Without objection, the Chair is authorized to declare recess at any time.

Before I deliver my opening remarks, I wanted to note that, today, the Committee is meeting both in person and virtually, which in my case is probably a good thing since I'm currently on day five of self-isolation due to mildly symptomatic COVID.

I want to announce a couple of reminders to the Members about the conduct of the hearing. First, Members and staff who were attending in person may choose to be masked, but it is not a requirement. However, any individuals with symptoms, a positive test, or exposure to someone with COVID–19 should wear a mask while present.

Members who are attending virtually should keep their video feed on as long as they are present in the hearing. Members are responsible for their own microphones. Please also keep your microphones muted unless you're speaking.

And finally, if Members have documents that they wish to submit for the record, please email them to the Committee Clerk, whose email address was circulated prior to the hearing.

Well, good morning, and welcome to our Members and witnesses. For today's hearing, I'm proud to announce that Representative Perlmutter and I, in our copious spare time, have been conducting experiments on a groundbreaking topic in nuclear physics. We are excited to share the results of that effort today as a preprint, and we plan to submit it to the *Reviews of Modern Physics.*

Well, just kidding, folks. What I'm actually referring to is an automated rip-off of a seminal paper published in the *Journal of Physical Review* in 1939 by Enrico Fermi called "Neutron Production and Absorption in Uranium," which has certain applications and relevance to nuclear power and nuclear weapons. We took Dr. Fermi's paper and ran it through a free online fake text generator that uses artificial intelligence (AI) to disguise plagiarism, and this took about 15 seconds. We then took about five minutes to tweak a few sentences to disguise the true source a little better. And once it was ready, we ran this paper through two well-respected plagiarism checkers. Each of these detectors found our fake paper was, and I quote, "100 percent unique and 0 percent plagiarism." Not surprisingly, these fake content generators have presumably been tuned up to generate low plagiarism scores, sort of the spambot equivalent of the generative adversarial network AI technique that's used to generate deep fake images and videos. And we even sent it over to the Inspector General at the National Science Foundation (NSF).

Now, any real human physicist peer reviewer for a journal or an NSF grant proposal would notice immediately that this paper uses silly technical jargon and plagiarizes from a very famous paper, and they would also find it unconventional that the report was authored by two sitting Congressmen and includes an acknowledgement to our Ranking Member Jay Obernolte. But you can imagine how bad actors might use tools as we did to sneak plagiarized content past journal editors and peer reviewers.

The AI-assisted plagiarism tool we used to make this fake paper is only one of the many in the arsenals of paper mills. These are

criminal enterprises that sell authorship credits for the fraudulent papers they place in academic journals. Scientific disciplines such as the life sciences, which rely heavily on images to communicate the results of experiments, are popular targets for fraud because of how easy it is to manipulate images.

Now, with the advent of sophisticated natural language processing software, it's becoming just as easy to churn out fraudulent but outwardly coherent text. Add in a few basic templates and the creation of hundreds of papers, complete with figures and citations, becomes the work of an afternoon. And much to the—this is much to the disgust of real scientists who might spend months on a single paper.

The scientific community must rise to meet this challenge, and it is already taking the first steps. Journals are looking for new ways to collaborate in detecting fraud during the review process. One recent effort is the STM (International Association of Scientific, Technical, and Medical Publishers) Integrity Hub, which would serve as a platform for journals to share dedicated fraud detection tools. The first tool under construction will flag the simultaneous submission of papers to multiple journals, which is a strong indicator of paper mill activity.

There's also a strong international community of researchers unaffiliated with publishers, many of them volunteers, who work to identify fraudulent papers following publication. Just as automation is enabling those committing fraud, it is also being used by these researchers to combat it. Next-generation plagiarism checkers don't just compare text to text, but intelligently scan for indicators of AI-generated text. Other tools detect manipulated images or identify erroneous science within the text of the paper itself. The automation arms race is upon us. We are here today to discuss how researchers and publishers can develop tools and policies that will help them stay ahead of the paper mills.

As we discuss scientific misconduct today, one of the most important things to keep in mind is the scale of the problem. Hundreds of papers with signs of fraud are indeed a serious concern. However, according to the NSF, a whopping 2.9 million papers were published last year alone. The number of cases of fraud must be viewed within that context. Creating and maintaining a body of scientific literature without flaws of any kind is an impossible quest, but published scientific literature remains the greatest body of human knowledge in the world, and it is our responsibility to look after its integrity.

This effort begins with a public dialog. As Dr. Fermi said famously, "Whatever nature has in store for mankind, unpleasant as it may be, men must accept, for ignorance is never better than knowledge."

I look forward to earning more knowledge today with the help of our esteemed witnesses.

[The prepared statement of Chairman Foster follows:]

Good morning and welcome to our Members and witnesses.

For today's hearing, I'm proud to announce that Representative Perlmutter and I have, in our spare time, been conducting experiments on a groundbreaking topic in nuclear physics. We're excited to share the results of that effort today as a preprint. We plan to submit it to *Reviews of Modern Physics.*

Just kidding. What I'm actually holding is a cheap rip-off of a seminal paper called "Neutron Production and Absorption in Uranium," which was published in the journal Physical Review in 1939. Its author was Enrico Fermi. We took Dr. Fermi's paper and ran it through a free online text generator that uses artificial intelligence to disguise plagiarism. This took 15 seconds. We then took five minutes to tweak a few sentences to disguise their true source a little better. Once it was ready, we ran this paper through two well-respected plagiarism checkers. We even sent it over to the Inspector General at the National Science Foundation. Each of these detectors found our fake paper was, and I quote—"100% unique, 0% plagiarism."

Now, any real physicist peer reviewer for a journal or an NSF grant would notice immediately that this paper uses silly technical jargon and plagiarizes from a very famous paper. They would also find it unconventional that the report was authored by two sitting Congressmen and includes an acknowledgement to Ranking Member Jay Obernolte. But you can imagine how bad actors might use tools as we did to sneak plagiarized content past journal editors and peer reviewers.

The AI-assisted plagiarism tool we used to make the fake paper is only one of many in the arsenals of "paper mills." These are criminal enterprises that sell authorship credits for the fraudulent papers they place in academic journals. Scientific disciplines such as the life sciences, which rely heavily on images to communicate the results of experiments, are popular targets for fraud because of how easy it is to manipulate images. Now, with the advent of sophisticated natural language processing software, it is becoming just as easy to churn out fraudulent but coherent text. Add in a few basic templates and the creation of hundreds of papers—complete with figures and citations—becomes the work of an afternoon, much to the disgust of real scientists who might spend months on a single paper.

The scientific community must rise to meet this challenge, and it is already taking the first steps. Journals are looking for new ways to collaborate in detecting fraud during the review process. One recent effort is the STM Integrity Hub, which will serve as a platform for journals to share dedicated fraud detection tools. The first tool under construction will flag the simultaneous submission of papers to multiple journals, a strong indicator of paper mill activity.

There is also a strong international community of researchers unaffiliated with publishers, many of them volunteers, who work to identify fraudulent papers following publication. Just as automation is enabling those committing fraud, it is also being used by these researchers to combat it. Next generation plagiarism checkers don't just compare text to text, but intelligently scan for indicators of AI-generated text. Other tools detect manipulated images or identify erroneous science within the text of the paper itself. The automation arms race is upon us. We are here today to discuss how researchers and publishers can develop tools and policies that will keep them ahead of the paper mills.

As we discuss scientific misconduct today, one of the most important things to keep in mind is the scale of this problem. Hundreds of papers with signs of fraud are indeed a serious concern. However, according to NSF, a whopping 2.9 million papers were published last year alone. The number of cases of fraud must be viewed within that context. Creating and maintaining a body of scientific literature without flaws of any kind is a quixotic quest.

But published scientific literature remains the greatest body of human knowledge about the world, and it is our responsibility to look after its integrity. This effort begins with a public dialogue. As Dr. Fermi said famously—

"Whatever Nature has in store for mankind, unpleasant as it may be, men must accept, for ignorance is never better than knowledge."

I look forward to earning some more knowledge today with the help of our esteemed witnesses.

I now yield to Ranking Member Obernolte for his opening statement.

Chairman FOSTER. And I now request unanimous consent to include in the record for this hearing both the real paper by Dr. Fermi and the fake one that we created, as well as a letter from the NSF Inspector General about they—how they tried to detect our sleight of hand.

Many thanks to Inspector General Allison Lerner, Dr. Aaron Manka, and their colleagues for their help in this.

And now the Chair will now recognize the Ranking Member for the Subcommittee on Investigations and Oversight, Mr. Obernolte, for an opening statement.

Mr. OBERNOLTE. Thank you very much, Chairman Foster, and I'm sure everyone here on the dais joins me in wishing you a speedy recovery and wishing that your illness remains asymptomatic. We're looking forward to having you back with us in person.

And I want to thank you for convening this hearing on an incredibly important topic, the topic of research integrity. It really is a topic that underpins our entire system of academic research here in this country.

A couple of years ago, I went back to graduate school to finish my doctorate, and finishing that dissertation was one of the hardest things I've ever done in my life. An important part of any research is to review the field of literature and the body of work on your research topic to determine exactly what's been done before and what the state-of-the-art in your research is. And I'll tell you, as I was going through the research in my field, it never occurred to me that some of those papers might be fraudulent.

That's one of the reasons why this hearing is so important is to call attention to what is a cutting-edge field in the body of academic research and literature, you know, this emergence of fraudulent research and paper mills and also to cast some light on the ways that technology can both enable this bad behavior by creating powerful tools that anyone could use to generate fraudulent papers, but also in combating the spread of fraudulent research by identifying the papers that might have been generated with artificial intelligence technology. So it's one of the reasons why I'm very much looking forward to this hearing.

I actually think that this is one of the things that we here in Congress can do very effectively, which is to simultaneously be a podium for the dissemination of information such as this because spreading awareness of this is going to be key to preventing the proliferation of this bad behavior. But I also think we have a role to play in catalyzing more research into how prevalent this problem is, in funding some Federal research into identifying the spread of the problem and identifying not only the causes of the problem but also some of the technology-based solutions to that problem, and in general raising awareness of this issue. And I also hope that everyone in the academic community joins us in recognizing just how destructive these paper mills have the potential to be for research integrity in general. I think the greater awareness that we have not only this is a problem but the—of the need to make severe penalties apparent for those who engage in destructive behavior like this, I think that's going to be key to controlling the spread of this problem.

So, Mr. Chairman, again, thank you very much for convening the hearing. I'm looking forward to hearing from our witnesses. I yield back.

[The prepared statement of Mr. Obernolte follows:]

Good morning. Thank you, Chairman Foster, for convening this hearing. And thanks to our witnesses for appearing before us today.

We are here today to discuss one of the most important aspects of scientific work, and the objective trust it instills, research integrity. As a member of the academic community myself, I am both troubled and perplexed to hear about the issue of paper mills. I am troubled because of the potential harm that these fraudulent papers can do to the scholarly record and perplexed by the motivations of researchers who choose to buy papers from a paper mill.

There is a saying in academia, referenced in today's Hearing Charter: "publish or perish". This mentality, along with other stringent career requirements internationally, seems to be driving some researchers to pad their resumes with paper mill papers.

Given our role in authorizing and overseeing the national research enterprise, I think it is important that we recognize this dynamic, while also thinking about how we can prevent this bad behavior. It is vital that publishers and universities remain diligent in preventing these fraudulent publications, and that there are consequences for engaging in this bad behavior.

As so often happens, the advancement of technology is an important tool to help the academic community rise to this challenge. Emerging technologies like AI are being used today to combat fraud by detecting plagiarism and faulty data. We should be wary though, because as these tools advance, so do tools to enable more bad behavior. One stark example is presented today by Chairman Foster's experiment -using an AI tool to create a fake academic paper. Even more problematic—this paper was not flagged as plagiarism by advanced plagiarism tools. Technology brings us new opportunities, but also new challenges. To combat this fraud it is important that the community remains diligent about the strengths and weaknesses of technology, and considers how additional investments in research can help to address this problem.

This is one area where I believe the Federal Government can play an important role—funding additional research on fraud detection. By placing emphasis and resources on research to create tools to help detect and flag fraudulent papers, federal research agencies can provide valuable input on what methods and tools should be considered best practices.

I am looking forward to hearing from our witnesses today. Each of them represents an important perspective in the academic community on how to combat this issue at a different stage in the process.

Thank you, Chairman Foster, for convening this hearing. And thanks again to our witnesses for appearing before us today. I look forward to our discussion.

I yield back the balance of my time.

Chairman FOSTER. Thank you.

[The prepared statement of Chairwoman Johnson follows:]

Good morning. Today's hearing will consider what seems to be a growing threat to the integrity of scientific publishing. The number of papers retracted in 2021 crossed 3,500, and volunteer sleuths find hundreds of cases of research misconduct each year.

I do not want to suggest that scientific journals are not paying attention to research misconduct. Quality control in paper submissions is a journal's bread and butter. Their reputations are a direct result of how successful they are in keeping fraudulent content out of print. But I also understand that if the goal is to keep 100% of fraud, fabrication, and plagiarism out of print, the odds are not in their favor.

With the dawn of foreign paper mills, the production of fraudulent content is now systematic. Language models powered by artificial intelligence are growing more sophisticated every day, making it easier than ever to produce fake content that looks authentic, or plagiarize real content so that it looks original.

As the methods of bad actors grow more powerful, we need to consider whether the scientific publishing enterprise is arming itself accordingly. Do journals have access to cost-effective, automated tools to assist with detecting misconduct before they even get to the peer review stage? Are there any automated tools that peer reviewers themselves can use to assist in their evaluation of original research? Are journals both motivated and equipped to investigate and adjudicate in a timely fashion any claims of misconduct that might be made about a paper that they have already published? Do journals always make it clear when it an article has been retracted for misconduct, so that the influence of the offending science is curtailed appropriately?

Our hearing today is focused on scientific journals, which are privately managed and funded. Prevention and detection of misconduct in federally funded research is its own critical issue. But I want to underscore that because of how scientists lean

on the other work of others, scientific integrity in privately-funded research is still a public good. Remember that in order to "see further" in his research, Sir Isaac Newton "stood on the shoulders of giants." Scientists use the published work of others to inform their own findings. Those other scientists are often halfway around the world, trying to publish and get ahead in an environment that the United States doesn't control.

If fraudulent work from any nation is allowed to persist in the scientific literature, it can undermine the good faith efforts of honest researchers. It can even influence laws or the behavior of the public to disastrous effect. Consider the fraudulent Wakefield paper, first published in 1998, which suggested childhood vaccines cause autism. A savvy journalist raised alarms about the critical flaws in this paper in 2004, but it was not officially retracted until 2010. It wreaked untold harm on public health in the interim.

I commend the volunteers like Dr. Byrne and Dr. Stell for their dedication to integrity in the scholarly record. The work that you and your peers do is a true service to the public. I know that it is often done at great personal sacrifice. I also want to commend Mr. Graf and STM for acknowledging the threats to your industry and for pursuing some scalable tools to address it. I look forward to hearing today about how government can be a partner to you going forward.

I yield back.

Chairman FOSTER. And at this time, I'd like to introduce our witnesses. Our first witness is Dr. Jennifer Byrne. Dr. Byrne is a Professor of Molecular Oncology at the University of Sydney. Her research has helped inform international debate on systemic fraud within the preclinical research literature, and she's an advocate for improved post-publication error reporting and correction. Dr. Byrne was a keynote speaker for both the 2021 Computational Research Integrity Conference and the Singapore Research Ethics Conference. She also chaired the Paper Mill Symposium at the 2022 World Conference on Research Integrity.

After Dr. Byrne is Mr. Charles Graf—Chris Graf, excuse me. Mr. Graf is the Research Integrity Director and Leader of the Editorial Excellence Team in Springer Nature. He also chairs the Governance Committee of the STM Integrity Hub, an initiative launched in early 2022 to help publishers collaborate to protect research integrity. Mr. Graf previously served as the Chair of the Committee on Publication Ethics, or COPE, and as a member of the Program Committee for the Seventh World Conference on Research Integrity.

Our final witness is Dr. Brandon Stell. Dr. Stell leads a research team with the French National Centre for Scientific Research, CNRS, studying the processing of sensory information in the brain. In 2012, Dr. Stell co-founded the website PubPeer.com to provide scientists with a forum to discuss the published research literature. PubPeer has since grown to be one of the leading sites for scientific discussion with a dedicated community of users who have helped strengthen the scientific record by exposing and correcting its weaknesses.

As our witnesses should know, each of you have five—whoops—yes, excuse me. As each of you should know, you will have five minutes for your spoken testimony. Your written testimony will be included in the record for the hearing, and you will all—when you all have completed your spoken testimony, we will begin with questions. Each Member will have five minutes to question the panel.

And we will start now with Dr. Byrne.

TESTIMONY OF DR. JENNIFER BYRNE,
DIRECTOR OF BIOBANKING,
NEW SOUTH WALES HEALTH PATHOLOGY;
PROFESSOR OF MOLECULAR ONCOLOGY,
UNIVERSITY OF SYDNEY

Dr. BYRNE. Thank you very much. So it's a pleasure to be here today. And I would very much like to thank Chairman Foster and Ranking Member Obernolte and all of the distinguished Members of the Committee. So my name is Jennifer Byrne, and I am a cancer researcher who has been studying what I believe to be systematic research fraud for about the past seven years.

As we have heard, paper mills are commercialized organizations that provide undeclared services to authors of scientific and scholarly publications including fabrication, fabricated data, and manuscripts. This represents a significant threat to science in terms of both its practice and its reputation. And the literature must take a no-tolerance approach toward papers that may have been constructed solely for career or commercial gain.

So there are a number of major factors that either drive authors toward paper mills or enable their activities, the first of which is unrealistic publication requirements that can be leveled across a broad range of authors, such as academic students and medical doctors, who may be—who may not be able to achieve the publication requirements that their institutions require of them.

But the—more in the field of publishing there has been an increasing focus upon a commercial focus through increasing use of author-paid publication services. These add to digital publishing capacities over the last 20 years that have greatly increased the capacity for papers to be published rapidly and also have enabled the creation of new journals. In contrast, the experimental sciences have not experienced the same capacity to increase their rate of data production.

Finally, a very important issue is the imbalance that currently exists between the production and the correction of scientific and academic publications. So most systems require appropriate balances between production and quality control. For example, cooking in a kitchen requires somebody to clean the kitchen afterwards. But at the moment in the scientific literature, the activity of production is greatly—it greatly overwhelms the capacity to clean and remove waste from the literature. This is a major advantage for research fraud and fraudulent publications because once they are published, they are very difficult to remove.

So the scope of the presence of fraudulent papers from paper mills within the scientific literature is largely unknown and—because this has been understudied. We recently screened just under 12,000 human gene research papers where we identified over 700 papers with errors that could signal paper mill involvement. Extrapolating from screening a very tiny fraction of the literature, I would estimate that the human gene literature contains more than 100,000 papers that have been produced by paper mills. This is a very serious issue, and the overall presence of paper mill publications and literature could be much higher because, obviously, many disciplines beyond human gene research have been targeted.

So the possible ramifications of large numbers of fraudulent papers are very concerning. For the research community, it is very likely that these papers are already misleading researchers in their research directions. They can damage research careers at all stages, encourage the support of unproductive research directions, and slow research translations through opportunity costs.

So clearly, given this scale of paper mill contributions, automated tools are necessary for the identification of the products of paper mills. We have used automation to screen papers for wrongly identified nucleotide sequences. These are reagents that are used in experiments, and their identities cannot be determined by the human eye, but they can be verified by appropriate detectors.

So the Seek & Blastn tool that was created by Dr. Cyril Labbe in Grenoble in 2017 uses an automated system of detection. Experience with this tool indicates that it provides a scale that cannot be matched by human experts. But its results need to be checked by humans in order to avoid false accusations of research errors, and clearly, this type of support can be difficult to obtain through research grants.

Tools such as Seek & Blastn can also be used by paper mills to remove errors from their papers and create papers that are more plausible and more likely to be published. So it is very important in my view that we move toward targeting paper mills through features that represent their business model as opposed to features of their products.

Publishers are now likely to be actively screening manuscripts for features of paper mills as an attempt to both detect and deter future submissions. We have proposed that another method that could be taken would be to require all research manuscripts to be posted to preprint service at the time of submission to reduce the duplicate submissions that Chairman Foster referred to in his opening address. We also believe that more aggressive steps are required to specifically disrupt the paper mill model such as to delay manuscript submissions through compulsory registrations at least one year prior to manuscript submission. This would not deter experimental scientists but would greatly damage the rapid publication timeframes that paper mills rely upon.

We would also like to see journals turning the same tools that they're using for screening manuscripts into their own archives to identify the papers from paper mills that have likely already been published and that are already misleading researchers in their daily work.

The Committee on Publication Ethics has recently described the need for retraction processes to rapidly adapt in response to the possibility of paper mills. We have proposed that journals could rapidly flag papers with verifiable errors using neutrally-worded notices such as editorial notes before investigation starts, as opposed to when they conclude so that researchers can be aware of papers having problematic features.

So, in summary, paper mills represent an unprecedented challenge to scientific and academic publishing, but they also provide a tremendous opportunity to enact transformational change. This can be achieved by increasing the oversight of scientific publishing, recalibrating our capacity to correct published information, as well

as to produce new information and overhauling the reward systems that underpin the careers of researchers and other professionals who publish it within the academic literature.

Thank you very much again for this opportunity to speak before the Committee, and I'll be very happy to answer any questions.

[The prepared statement of Dr. Byrne follows:]

Written testimony of
Jennifer Anne Byrne PhD
Conjoint Professor of Molecular Oncology
Faculty of Medicine and Health
University of Sydney, Australia
Before the House Committee on Science, Space, and Technology

For the hearing titled:

"Paper Mills and Research Misconduct: Facing the Challenges of Scientific Publishing"

July 20, 2022

Representative Foster, and all distinguished members of the committee,

Thank you for the invitation to join the hearing today. I am a conjoint Professor in the Faculty of Medicine and Health at the University of Sydney, Australia, where my area of expertise is the analysis of human genes in cancer and publication integrity. I am honoured to join the committee to discuss current and future challenges in securing the scientific literature from fraudulent academic papers.

Introduction

As Isaac Newton famously stated, scientists see further by standing on the shoulders of giants. When these shoulders are built from unreliable or fraudulent research, progress stalls and the funds that support our best and brightest researchers are wasted.

Paper mills are commercial organisations that allegedly provide undeclared services to authors of scientific and scholarly publications, including fabricated data and manuscripts (1). The threat of paper mills to scientific publishing and integrity has no parallel over my 30-year scientific career. The systematic production of large numbers of fraudulent or fabricated manuscripts harms science, both in its practice and reputation. The scientific literature must take a no-tolerance approach towards papers that may have been constructed solely for career or commercial gain.

Before answering the questions posed by the House Committee, I would like to outline 3 factors that I believe are important in driving the use of paper mills:

(1) Unrealistic publication requirements

Most scientists and academics experience expectations to produce research articles. Pressure to publish becomes problematic when institutions impose publication requirements to either retain academic positions or meet career goals. In some cases, genuine research efforts from academics, students and medical doctors may be insufficient to achieve these requirements. Unrealistic publication quotas that are applied over populations can create large markets for paper mills.

(2) The increasingly commercial focus within scientific and academic publishing

Over the past 20 years, scientific publishing has changed in response to digital publishing and the growing use of author publication fees to generate publisher income. Author-paid publishing has led to a more profit-focussed environment that can reward publication quantity over quality. Digital publishing similarly allows for the publication of more articles, and the creation of new journals that could increasingly compete for the same pools of manuscripts. This is a major issue for scientific fields that produce data through experiments. Research funding has showed limited growth over the past 50 years, and many types of experiments remain difficult, expensive and slow. *While online digital publishing can rapidly expand in response to perceived market forces, the production of experimental results cannot expand with the same speed.* In contrast, paper mills can supply fabricated manuscripts to journals in rates and numbers that experimental scientists cannot produce.

(3) Imbalance between the production and correction of scientific and academic publications

Most systems require an appropriate balance between production and quality control. For example, hospitals rely upon quality cleaning services, and communities require regular waste collection. *Astonishingly, the activities of cleaning and waste removal are largely missing from scientific and academic publishing.* Research funding overwhelmingly supports new knowledge production, and there are comparatively few funds devoted to detecting and correcting published errors. Similarly, scientific and academic publishing focusses upon the publication of new manuscripts, and not on correcting or removing unreliable information. *This imbalance between production and correction means that once fraudulent research is published, it is very difficult to remove from the literature.*

The scope and the ramifications of the presence of fraudulent publications from paper mills within the scientific literature

Key takeaway: In the field of human gene science alone, the number of potentially fraudulent articles could exceed 100,000 original papers. Research is urgently needed to illuminate the history and trajectory of paper mill contributions, so that effective actions can be designed and implemented.

The full **scope** of paper mill contributions to the scientific literature is poorly understood, and likely to be underestimated. We have recently screened just under 12,000 human gene research papers for gene sequence errors that may be associated with paper mill support (2). We identified over 700 papers with errors that could signal paper mill involvement (2). Based on this proportion, the number of paper mill contributions to the human gene literature could exceed 100,000 original articles. This estimate may seem shocking but is likely to be conservative, as (i) paper mills may have been contributing to the human gene research literature for over 15 years across many individual journals, (ii) not all paper mill-supported papers contain the same types of errors, and (iii) some papers could be error-free. As paper mills have also been alleged to have targeted other topics, the total number of paper mill supported publications could be several fold higher. This is supported by research from the Committee on Publication Ethics, where interviewed journals estimated that 2-46% of manuscript submissions could come from paper mills (1).

Given this predicted scale, *research is urgently needed to define the scope of paper mill contributions to the academic literature*. This requires research to identify features of paper mill manuscripts in different fields, and to develop, improve and apply automated detection tools at scale. We need this research to understand how many paper mill-supported articles

have been published, how these articles may have changed over time, and in what direction paper mill manuscripts may be continuing to evolve.

The possible **ramifications** of large number of fraudulent papers from paper mills are very concerning. Human gene research papers with features of possible paper mill involvement are cited by the preclinical and clinical literature (2). This suggests that paper mill articles are *misleading researchers and research directions*. Pursuing fabricated gene research could damage biomedical research careers at all stages, encourage the support of unproductive research directions, slow clinical and industry research translation through opportunity costs, and reduce confidence in research and the scientific method. The challenge of distinguishing genuine and fraudulent publications may cause researchers to abandon particular research fields. Paper mill articles could therefore decrease publication outputs from genuine research, compounding their damaging effects on scientific progress.

How we have used automation to identify fraud, and how we anticipate automation being used in the future to either combat fraud or to perpetuate it

Key takeaway: Due to the estimated scale of paper mill contributions, automated tools are necessary for their identification. More resources are required to support the development, testing and application of automated tools.

We have used automation to screen publications for wrongly identified nucleotide sequences, or incorrect gene sequences (2). Nucleotide sequences are used in experiments that study genes from humans and other organisms. They are like barcodes, in that they convey a meaning that cannot be read by humans, but can be verified by an appropriate detector. Incorrect nucleotide sequences can signal possible research fraud where their verified identities could not have produced the results that papers describe.

The Seek & Blastn tool created by Dr Cyril Labbé in Grenoble, France, uses an automated system of detection. Our experience with this tool over the past 5 years is that it provides scale that cannot be matched by human experts (2). Other automated tools are available to detect different questionable features of publications. However, the lack of research investments in error detection and correction mean that many automated tools, including Seek & Blastn, have not been developed or applied to their full potential. The results of automated tools also need to be checked by human experts, who require salary support that is difficult to obtain through research grants. Research on publication error detection needs to recognise the need for human experts and to provide training and career pathways that lead to rewarding careers in the field of error detection.

At least some paper mills are likely to use automation to produce research manuscripts at scale. Just as freely accessible tools such as Seek & Blastn can help researchers to identify unreliable research results, these tools could also be used by paper mills to produce more plausible manuscripts. The capacity to artificially generate highly plausible versions of experimental results could render paper mill manuscripts more difficult to detect. These developments *highlight the urgent need to deter paper mill submissions by targeting specific requirements of the paper mill business model,* as opposed to manuscript features that may not apply to all disciplines. This will be discussed further below.

How publishers and authors have addressed, or failed to address, the detection of research misconduct in their articles

Key takeaway: There are currently few incentives for publishers to correct problematic research, or for other scientists to report it. We urgently need to increase capacity to achieve timely corrections to the published record.

Major publishers and journals are now focussing on detecting manuscripts from paper mills and attempting to deter future submissions (1). However, variable screening approaches across different journals and publishers can mean that previously rejected manuscripts from paper mills can be accepted elsewhere. *Uniform screening practices and universal requirements to post all research manuscripts to preprint servers at the time of submission could reduce submissions from paper mills.*

Some publishers and journals have also implemented new manuscript standards that are intended to deter paper mill submissions. However, some suggested improvements can be very easily accommodated by paper mills and may not serve as useful deterrents. *More aggressive steps are required that will specifically disrupt the paper mill business model.* One such approach would be to delay manuscript submissions through a compulsory registration process at least one year prior to manuscript submission. This requirement could be designed to be compatible with the timeframes of genuine research, while seriously disadvantaging the rapid publication timeframes that are likely to be valued by paper mills.

While many publishers are now screening for paper mill features in manuscripts, few journals appear to be applying the same screening methodologies to their published archives. *Tools that can be used to screen manuscripts for features associated with paper mills should also be applied to published articles.* There are currently few incentives for journals to proactively screen their archives for erroneous publications and to then instigate retractions. Incorporating measures of post-publication correction into metrics such as the journal impact factor could incentivise post-publication corrections and tangibly reward proactive journals and publishers.

I can speak my experiences in bringing gene research publications with errors to the attention of journals and publishers. While a small number of journals have been highly supportive, many other journals appear to ignore error descriptions and requests for publications to be investigated. These experiences have been described by other teams, and collectively discourage error reporting.

The Committee on Publication Ethics (COPE) has recently described the need for retraction processes to adapt to publications from paper mills (1). We have proposed that journals could rapidly flag papers with verifiable errors (such as wrongly identified gene sequences) using neutrally worded notices before journal investigations start. These notices could be published very quickly in response to error notifications and could be transferred between different platforms, including PubPeer. A more rapid and responsive post-publication correction system would also encourage more researchers to recognise and report errors within the literature.

Summary

Paper mills represent an unprecedent challenge to scientific and academic publishing, but also provide an opportunity to enact transformational change. This can be achieved by increasing the oversight of scientific publishing, recalibrating our capacity to correct

published information, and changing the reward systems that underpin the careers of researchers and other professionals who publish within the academic literature. We must now commit to building and empowering the human and infrastructure capabilities that will be required to repair and safeguard our scientific literature for future generations.

References

(1) COPE. and STM. (2022) Paper Mills - research report from COPE & STM - English. https://doi.org/10.24318/jtbG8IHL.

(2) Park, Y., West, R.A., Pathmendra, P., Favier, B., Stoeger, T., Capes-Davis, A., Cabanac, G., Labbé,C. and Byrne, J.A. (2022) Identification of human gene research articles with wrongly identified nucleotide sequences. Life Sci. Alliance, 5, e202101203.

Professor Jennifer Byrne
Career summary

- BSc (Hons 1, University Medal) (1988) PhD (1993), University of Queensland, Australia
- NHMRC CJ Martin Postdoctoral Fellow (1993-1997) (France, Australia)
- Group Leader (1998-2019), Deputy Unit Head (2004-2008), Unit Head (2009-2019), Children's Cancer Research Unit, The Children's Hospital at Westmead, Australia
- Deputy Director, Kids Cancer Alliance Translational Cancer Research Centre (2018-2019), Australia
- Conjoint Senior Lecturer (2003-2006), Conjoint Associate Professor (2007-2016), Academic Leader (2017-2019), Conjoint Professor (since 2017), The University of Sydney, Australia
- Director of Biobanking (since 2019), NSW Health Pathology, Australia

Contributions to field of research: Byrne first reported the existence of incorrectly identified nucleotide sequence reagents within pre-clinical cancer research publications. She recognised that nucleotide sequence reagents represent a class of verifiable reagent that are prone to acquiring errors. These insights, combined with the descriptions of nucleotide sequence reagents in hundreds of thousands of research publications, underpinned the creation of the first semi-automated tool Seek & Blastn to fact-check the published identities of nucleotide sequence reagents. This fact-checking capacity had been present in the biomedical literature for decades but had not been previously recognised or leveraged. Byrne has leveraged features of papers with wrongly identified sequences that they have discovered to inform international debate on the possibility of systematic research fraud within the pre-clinical research literature, and to advocate for improved post-publication error reporting and correction. Seek & Blastn is now used to screen manuscripts at multiple biomedical journals as well as COVID-19 preprints through the international ScreenIT Group.

International and national profile: Byrne is known for her research towards understanding human gene functions, cancer genetics, cancer predisposition in children, improving biobank operations and support of biomedical and health research, and error detection and correction within the biomedical literature. Byrne was included as one of Nature's 10 people who mattered in 2017 for her error detection research, which has also been highlighted by Nature News (2017, 2020, 2021 (twice)), Retraction Watch (2017 (twice), 2018, 2019, 2021), The Atlantic (2018), Undark Magazine (2018), Wall Street Journal (2020), The Scientist (2021), and Times Higher Education (2021). She wrote about the need for clearer scientific communication in The Conversation in 2018. Recent international speaking invitations include as a keynote speaker and panellist at the CRI-CONF Computational Research Integrity Conference (2021), and as an invited/ keynote speaker at the Singapore Research Ethics conference (2021), the Science Integrity Symposium in Germany (2022) and the Science Studies Colloquium, Denmark (2022). Byrne chaired the paper mill symposium at the 2022 World Conference on Research Integrity.

Research support: AUD$15.7 million in funding as a chief investigator in the last 5 years, including National Health and Medical Research Council Ideas Grant APP1184263 as CIA "Prevalence and impact of fraudulent cancer research publications targeting the functions of human genes", AUD$4M from the NSW Luminesce Alliance to support paediatric cancer predisposition screening (2019-2022), AUD$1M from Frontier Health Medical Research to support the development of phage therapy in Australia (2021-2022)

Supervision and mentoring: Principal supervisor: 2 postdoctoral fellows, 11 PhD students, 10 Masters and Honours students, one current PhD student. Deputy Postgraduate Co-ordinator (2006-2011), multiple awards from the University of Sydney for outstanding postgraduate student teaching and supervision (2003, 2005 (two awards), 2011). Byrne has mentored candidates applying for level E promotion at the University of Sydney since 2019.

Current professional involvement (selected)

- Member, Steering Committee, Brain Cancer Biobanking Australia, since 2014
- Member, Victoria Cancer Biobank Scientific Advisory Board, since 2019
- Chair, Scientific Advisory Group, NSW Health Statewide Biobank, since 2019
- Member, ScreenIT Group, since 2020
- Asia Pacific Research Integrity Network Meeting Program Planning Board, since 2020
- Member, Education and Training Committee, International Society for Biological and Environmental Repositories, since 2021
- Member, Australian Brain Cancer Mission Strategic Advisory Group, since 2022
- Board member, Association for Interdisciplinary Meta-research & Open Science, since 2022

Journal editorial boards

- Subject Editor, International Journal of Biological Markers, since 2014
- Editor-in-Chief, Biomarker Insights (2019-2021)

Most relevant publications (from most recent)

1. Schulz, R., Barnett, A., Bernard, R., Brown, N.J.L., **Byrne, J.A.**, Eckmann, P., Kilicoglu, H., Prager, E.M., Salholz-Hillel, M., ter Riet, G., Vines, T., Vorland, C.J., Zhuang, H., Bandrowski, A., Weissgerber, T.L. (2022). Is the future of peer review automated? BMC Research Notes **15**: 1-5.
2. Park Y., West R.A., Pathmendra, P., Favier, B., Stoeger, T., Capes-Davis, A., Cabanac, G., Labbé, C., **Byrne, J.A.** (2022). Identification of human gene research articles with wrongly identified nucleotide sequences. Life Science Alliance. **5**: e202101203.
3. Parker, L., **Byrne, J.A.**, Goldwater, M., Enfield, N.J. (2021). Misinformation: an empirical study with scientists and communicators during the COVID-19 pandemic. BMJ Open Science **5:** e100188.
4. Weissgerber, T., Riedel, N., Kilicoglu, H., Labbé, C., Eckmann, P., ter Riet, G., **Byrne, J.**, Cabanac, G., Capes-Davis, A., Favier, B., Saladi, S., Grabitz, P., Bannach-Brown, A., Schulz, R., McCann, S., Bernard, R., Bandrowski, A. (2021). Automated screening of COVID-19 preprints: Can we help authors to improve transparency & reproducibility? Nature Med. **27:** 6-7.
5. **Byrne, J.A.,** Park, Y., West, R.A., Capes-Davis, A., Cabanac, G., Labbé, C. (2021). The thin ret(raction) line: biomedical journal responses to reports of incorrect non-targeting nucleotide sequence reagents in human gene knockdown publications. Scientometrics. **126**: 3513-3534.
6. **Byrne, J.A.,** Christopher, J. (2020). Digital magic, or the dark arts of the 21st century- how can journals and peer reviewers detect manuscripts and publications from paper mills? FEBS Lett. **594**: 583-89.
7. Labbé, C., Cabanac, G., West, R.A., Gautier, T., Favier, B., **Byrne, J.A.** (2020). Flagging errors in biomedical papers: to what extent does the leading publication format impede automatic error detection? Scientometrics. **124**:1139-1156.
8. Labbé, C., Grima, N., Gautier, T., Favier, B., **Byrne, J.A.** (2019). Semi-automated fact-checking of nucleotide sequence reagents in biomedical research publications: the Seek & Blastn tool. PLOS ONE. **14**: e0213266.
9. **Byrne, J**. (2019). We need to talk about systematic fraud. Nature. **566**: 9-10.
10. **Byrne J.A.**, Grima N, Capes-Davis A, Labbé C (2019). The possibility of systematic research fraud targeting under-studied human genes: causes, consequences and potential solutions. Biomarker Insights. **14:** 1-12.
11. **Byrne, J.A.,** Labbé, C. (2017). Striking similarities between publications from China describing single gene knockdown experiments in human cancer cell lines. Scientometrics. **110**: 1471-93.
12. **Byrne J.A.** (2016). Improving the peer review of narrative literature reviews. Res Integrity Peer Rev. **1**: 12.

Chairman FOSTER. Thank you. And next is Mr. Graf.

TESTIMONY OF MR. CHRIS GRAF,
RESEARCH INTEGRITY DIRECTOR, SPRINGER NATURE;
CHAIR OF THE GOVERNANCE BOARD,
STM ASSOCIATION INTEGRITY HUB

Mr. GRAF. Thank you, Chairman Foster and Ranking Member Obernolte and esteemed Members of the Committee for inviting me. I'll give an overview following Jenny Byrne's overview of what research publishers are doing to safeguard research integrity.

As introduced previously, I'm Chris Graf, the Research Integrity Director at Springer Nature, which is one of the world's leading research publishers and Chair of the Governance Committee for the Integrity Hub, which is a collaborative tech initiative from the STM Association, which is the global trade body for research publishers.

My written testimony explains how the research publishing sector is facing one of its current challenges, namely that of paper mills and research misconduct. And I conclude in that written testimony that the opportunities exploited by paper mills are created somewhere upstream where research is done. Publishers can and are doing more to stop papers from paper mills, and other actors also have a responsibility, including the organizations that fund and employ and set policy for researchers and for research. I'd argue that's where the solution lies in a broad coalition of those who are able to act to make change happen.

I would like to use the rest of my five minutes for some background. I'll talk briefly about science then briefly about publishing, and then a little more extensively about paper mills. First, science, so trust in science remains strong. The 2021 survey from NORC at the University of Chicago reports that 48 percent of Americans have a great deal of confidence in the scientific community. And members of those scientific communities published—well, my numbers say 5 million peer-reviewed scientific articles in '21. That's from a dimensions data base. That contrasts with the 2.9 million that the NSF reported, so lots of millions, but somewhere between 2.9 and five. Within that, if we zoom in a little on COVID, there are now 630,000 COVID papers in the World Health Organization data base for literature on COVID, on coronavirus disease. Half of those were published last year in 2021. So that's a view across science and a bit about publishing that science.

But we're here to talk about paper mills and misconduct. So that's when things go wrong or actually very wrong. What happens then? Well, you may know but we retract scientific papers when un-addressable concerns are identified. Those concerns range from honest and fundamental errors that might be embarrassing for a researcher, but, you know, that researcher should be applauded for addressing them and for retracting those papers and clearing up the inaccurate information they've published. And they range from those honest and fundamental errors through questionable and misleading research, which could be naive and might be negligent, but probably isn't malicious, right through to misconduct, including that promoted by paper mills.

And that kind of retraction doesn't happen often. Historically, 4 in 10,000, peer-reviewed science articles are retracted after publication. Zooming in to look at that through the lens of COVID, about 300 of the 630,000 COVID papers published so far have been retracted. And I think that's similar to the general rate that I described earlier, 4 in 10,000.

You know, I think that's an indicator of significant and successful investments made into quality and into integrity by researchers first and also by publishers when it comes to publishing them. I'd argue that the contribution that research publishers make to quality and integrity is—well, I'd argue it's true. It's based on years of collaborative efforts. Publishers with other stakeholders for years have been developing and sharing resources about how to manage honest but fundamental mistakes through to the other end of the range that we talked about earlier, to misconduct and systemic—systematic manipulations.

And publishers continue to invest in screening for integrity, including routine checks for plagiarism that we've heard about, but they're being enhanced and improved, as well as other indicators for ethics and quality like the disclosures of conflicts of interest. Some publishers, as Jenny referred to, are beginning to roll out screening for image manipulation, which is a newer fingerprint that might indicate the presence of a paper mill. And both of these require investments not only in technology but in actual people to use that technology.

Even so, I agree that paper mills are a growing threat. Evidence suggests they're operating with relative freedom. When they find a way into a journal that has weak defenses, they certainly exploit that, and they do cause real damage. We've referred to that already. They steal credentials from legitimate researchers, for example, they con their way into editorial positions of power at journals, and then they use that to their advantage. And that's identity theft, and that's fraud. They do many other inappropriate things as well.

So let me close with what I think the challenge is. Legitimate researchers currently benefit from a largely trust-based system. Solving the paper mill problem without making publishing harder and less trust-based for the vast majority of legitimate researchers I think is the challenge. The new STM Integrity Hub shows how publishers are taking collective action and using their combined knowledge and technology to do just that.

So that's where I'll end. Thank you for the opportunity to present on how the publishing sector is responding to the challenges of research misconduct. It really was a privilege and is a privilege to be part of today's hearing. I genuinely look forward to your questions and also to continue to serve in any way that might be useful for you. Thank you.

[The prepared statement of Mr. Graf follows:]

SPRINGER NATURE GROUP

Springer
Science+Business Media, LLC
One New York Plaza, Suite 4600
New York, NY 10004-1562, USA

T +1 212 726 9200
www.springernature.com

Witness Biography and Written Testimony for the House Committee on Science, Space, and Technology, Subcommittee on Investigations & Oversight, hearing:

Paper Mills and Research Misconduct: Facing the Challenges of Scientific Publishing

Witness: Chris Graf, Chair, Governance Committee, STM Association Integrity Hub, and Research Integrity Director, Springer Nature
Dated: Monday July 18, 2022

Contents

Biography **2**

Written Testimony **3**

Introduction: Some things are better done together, research integrity is one 3
Let me start by presenting background 3
Challenges to research integrity: Errors, misconduct and fraud 4
Scale of fraudulent publications from paper mills 5
Continued investment in publishing technology is part of the answer 6
Action is needed beyond the publishing sector 7
Further reading 7

SPRINGER NATURE GROUP

Biography

Chris Graf is the Research Integrity Director (a position he has held since September 2021) and Leader of the Editorial Excellence Team (since January 2022) at Springer Nature. In these roles, he is in charge of leading the continued development and implementation of Springer Nature's research integrity strategy, leading a 20-person team accountable for upholding the highest standards for research integrity in all of Springer Nature's publications. Chris also serves the publishing sector with a voluntary role at the STM Association, the global trade association for academic and professional publishers, where he chairs a committee of senior publishing industry executives overseeing governance of the STM Integrity Hub. The STM Integrity Hub is an initiative launched early 2022 to equip the publishers (and the scholarly communication community) with data, intelligence, and technology to protect research integrity.

Chris has decades of experience in the intersection of scientific publications, academic research, and the ethics of the scientific peer review process. He was a long-time volunteer and is a continuing advocate for the Committee on Publication Ethics ("COPE"), serving in a number of positions from 2005 to 2020 including treasurer, vice-chair, and co-chair, and has worked in a number of other positions at the intersection of research integrity and publishing. He has served on the World Conference of Research Integrity programme committee since 2017. He has written widely on the topic of publishing and academic integrity in the peer review process, including published articles on minimum report standards for life scientists and on how different models of peer review impact final research quality.

Chris started at Wiley in 2004, serving as a Publisher and then an Associate Editorial Director in clinical medicine. He published Wiley's first edition best practice guidelines on publication ethics. In 2011, he was promoted to serve on secondment as the Editorial Director for Health and Life Sciences in Australia. On his return to the UK in 2014 he worked in Wiley's innovation team as the Digital Learning Director and Learned Society Partnership Director, from 2014-2017. His time at Wiley continued with him becoming the Director of Research Integrity in the Wiley Open Research team, a post he held from 2017 to 2021. Chris began his career as an Editor for various publications and holds a bachelor's degree in Chemistry.

SPRINGER NATURE GROUP

Written Testimony

Introduction: Some things are better done together, research integrity is one

Thank you, Chairman Foster, Ranking Member Obernolte and esteemed members of the committee for inviting me to testify today and the opportunity to share with you an overview of the work we do to safeguard research integrity.

My name is Chris Graf. I am the Research Integrity Director at Springer Nature and also serve as the Chair of the Governance Committee for the STM Association Integrity Hub.

The written testimony below sets out how the research publishing sector is facing one of its current challenges, namely the challenge of paper mills and research misconduct. The testimony concludes that the opportunities exploited by paper mills are created 'upstream' where research is done. While publishers can and are doing more to stop papers generated by paper mills, other actors also have a responsibility, including the organizations that fund and employ researchers. That's where the solution lies, in a broad coalition of those who are able to make change happen. Some things are better done together, research integrity is one.

Let me start by presenting background

- **Trust in science remains strong**. 48% of Americans have a great deal of confidence in the scientific community per NORC's 2021 survey (General Societal Survey, *NORC at the University of Chicago*). The NORC survey and another from Pew both suggest this may be trending downwards since COVID, and could depend on the respondent's political alignment (*Pew Research Center*). Outside the US measures of public trust in scientists and professors are similarly high at +70% net trust, compared with nurses who score +91% and advertising executives who score −59% (UK 2021 Trust in professions survey, *IPSOS MORI*).
- **5,000,000 peer reviewed scientific papers** were published in 2021 (*Dimensions*).
- **630,000 COVID papers** are in the WHO database for global literature on coronavirus disease (*World Health Organization*).
- **A scientific paper is retracted when serious and unaddressable concerns are identified** about the reliability of the scientific content presented in a paper. These concerns range from honest but fundamental errors (Retractions: A clean slate, *Nature*), via questionable or misleading research practices, to misconduct including that promoted by 'paper mills'.

- **Only 4 in 10,000 peer reviewed scientific articles are retracted** after publication. Sometimes this happens months later, sometimes years (Rethinking retractions, *Science*). We consider this to be an indicator of significant and successful investments in quality made by researchers and by publishers.
- **About 300 COVID papers have been retracted** (similar to the general rate of about 4 in 10,000 cited above) (*Retraction Watch Database*).
- **Publishers, with other stakeholders, have been developing and sharing resources** about managing honest but fundamental mistakes, as well as questionable practices, misconduct, and systematic manipulations for some years (Systematic manipulation of the publication process, *COPE*).
- **Publishers continue to invest in screening for integrity**, including routine checks for plagiarism as well as for other indicators of ethics and quality (like disclosures of conflicts of interest, description of ethics committee approval and funding sources). Both require investments in a mixture of technology and operations. Some publishers are beginning to roll out screening for newer concerns that indicate paper mills, like image manipulation (The fight against fake paper factories, *Nature*).
- **Paper mills are a relatively new and growing threat** in the research and publishing community. Evidence suggests they are operating with relative freedom. (Paper mills Research Report, *COPE/STM Association*).
- **Paper mills cause real damage.** (Tackling paper mills and bogus research: Some things are better done together, *Springer Nature*).
- **Legitimate researchers benefit from a largely trust-based system.** Solving the paper mill problem without making publishing harder or less trust-based for the vast majority of legitimate researchers is the challenge. (STM Integrity Hub, *STM Association*).

Challenges to research integrity: Errors, misconduct and fraud

Paper mills are fraudulent organizations that profit by systematically and duplicitously manipulating the systems and processes used to write and publish science. The causes of, and solutions to, the harm to the scientific record caused by paper mills are being discussed by stakeholders across the research publishing sector. Solving the paper mill problem without making publishing harder or less trust-based for the vast majority of legitimate researchers is the challenge.

Research publishers represented by the STM Association (the global trade association for academic and professional publishers) are concerned by the increase in research integrity issues in general, and particularly with paper mills. Paper mills are using increasingly advanced technologies to fabricate, plagiarize, and manipulate text, images, and research data. Paper

mills operate systematically and at scale, but misconduct by individual researchers (for example image manipulation or data fabrication) is another concern. Unwitting errors by researchers likewise pose a problem for the integrity of the scholarly record, but in and of themselves are not malicious (these are sometimes referred to as questionable research practices). Each of these areas of concern potentially undermines trust in science. They place a burden on publishers, editors, and peer reviewers. The STM Association and many of its members are concerned about the dynamics that drive these behaviors.

Scale of fraudulent publications from paper mills

The STM Association and COPE (Committee on Publication Ethics, a UK charity that brings together all those involved in scholarly research and its publication to strengthen integrity) commissioned a study of data submitted from a variety of leading publishers to get a sense of the scale of the problem. This study was published as a Research Report in June 2022 (https://publicationethics.org/files/paper-mills-cope-stm-research-report.pdf). The results show that the submission of suspected fake research papers with fake authorship is growing and threatens to overwhelm the editorial processes of some journals.

As part of the STM/COPE study, data on over 53,000 papers was analyzed. This data was shared by six publishers and spans a wide range of research disciplines. Overall the percentage of suspect papers being submitted to the journals that were investigated ranges from 2 to 46% (please note that these journals do not constitute a representative sample, and the findings are not generalisable). The analysis shows that most journals included in the analysis saw 2% suspected fake papers submitted and then, for journals where paper mills have been successful in getting papers accepted, those journals see a sharp increase in suspect submissions.

The Research Report identifies two areas of work publishers are currently undertaking:

- Pre-publication submission review: Tools and processes are increasingly being used to identify suspect papers early in the submission process.
- Post-publication review and retraction: There are a number of ways that a journal can identify a suspect paper already published, and standard processes to follow thereafter.

Interviews with a range of stakeholders including publishers, research investigators and journalists (including at *Retraction Watch*) show what we would characterize as a shared level of concern. There is a realization that all stakeholders need to work together to find long-term solutions. The Research Report ends with a call for collective action, and makes 5 Key Recommendations:

1. Engagement with institutions and funders
2. Continued investment in tools to detect potential paper mill papers
3. Educational exercise for editors and editorial staff
4. Investigation of protocols to impede paper mills

5. Review of the retraction process to take account of the unique features of papermill papers

Continued investment in publishing technology is part of the answer

Publishers, with other stakeholders, have been developing and sharing resources about managing honest but fundamental mistakes, as well as questionable practices, misconduct, and systematic manipulations for some years https://publicationethics.org/resources/flowcharts/systematic-manipulation-publication-process.

The STM Association and many of its members believe that reliable, state-of-the-art technology solutions can play an important role in addressing these challenges by flagging possible concerns before publication to support the editorial decision-making process. This allows publishers to prevent this material from entering the scholarly record. For this reason, in early 2022 STM launched the STM Integrity Hub https://www.stm-assoc.org/stm-integrity-hub/. The mission of the Hub is to equip the publishers and the scholarly communication community with data, intelligence, and technology to protect research integrity. The Hub will do that in three ways:

1. Fostering the sharing of intelligence and knowledge: For example, sharing experiences that publishers have in identifying papers that are produced by paper mills.
2. Building policies and frameworks: Everything we do in terms of safeguarding research integrity has to be firmly embedded in legal and policy frameworks, where we work closely with editorial and legal teams, as well as with organizations like COPE.
3. Building a platform that allows publishers to provide content to detect patterns across publishers in a safe and confidential way, but also to easily integrate third-party screening tools in their workflows that focus on specific questionable research practices, e.g. image manipulation.

Publishers and publishing technology providers have invested in systems to detect potentially problematic cases for years. For example, publishers and journals routinely check for plagiarism using investments they make in the publisher-initiated non-profit organization CrossRef, with the technology provider iThenticate, and integrated with most editorial software systems provided by publishing technology companies and paid for by publishers. Some larger publishers have also started to invest in development of their own technologies and operations to counter new challenges, for example challenges in image manipulation. However, the new and changing techniques adopted by paper mills require ongoing investments, and the significant financial burden is challenging to larger publishers and prohibitive to smaller ones. Additionally, working as individual publishers does not facilitate knowledge sharing. This is why the collaboration under the STM Integrity Hub is critical. It creates the organization needed for collective investment and effective action, and makes the burdens manageable.

Currently, over 20 publishers are actively participating in the Integrity Hub including commercial publishers, learned societies and university presses, representing all academic disciplines and both smaller, mid-sized and larger organizations. This reflects the collaborative nature of the initiative, and the strong commitment to maintain the integrity of science.

Action is needed beyond the publishing sector

To close, a final observation. The opportunities exploited by paper mills are created 'upstream' where research is done. While publishers can and are doing more to stop papers generated by paper mills, other actors also have a responsibility. Included among these other actors are the organizations that fund and employ researchers, who create research incentives and environments. The first Key Recommendation from the STM/COPE Research Report is to engage with institutions and funders, and so take a system-approach to the paper mill problem. That's where the solution lies. Some things are better done together, research integrity is one.

Further reading

Tackling paper mills and bogus research: Some things are better done together, *Springer Nature*, 28 April, 2022

Want research integrity? Stop the blame game, *Nature,* World View, 24 November 2021

What a massive database of retracted papers reveals about science publishing's 'death penalty', *Science*, 25 October 2018

Research integrity is much more than misconduct, *Nature,* Editorial, 3 June 2019

Systematic manipulation of the publication process, *COPE*, February 2019 updated December 2021

We need to talk about systematic fraud, *Nature,* Comment, "Software that uncovers suspicious papers will do little for a community that does not confront organized research fraud", says Jennifer Byrne, 6 February 2019

Swedish research misconduct agency swamped with cases in first year, *Nature,* news article, 13 September 2021

The fight against fake-paper factories that churn out sham science, *Nature,* news feature, 23 March 2021

China's clampdown on fake-paper factories picks up speed, *Nature,* news article, 1 October 2021

How to investigate allegations of research misconduct: a checklist, *Retraction Watch*, 8 January 2019

Revealed: The inner workings of a paper mill, *Retraction Watch*, 20 December 2021

How to find evidence of paper mills using peer review comments, *Retraction Watch*, 21 February 2022

20 ways to spot the work of paper mills, *Retraction Watch*, 9 February 2021

-- Ends --

SPRINGER NATURE GROUP

Biography

Chris Graf is the Research Integrity Director (a position he has held since September 2021) and Leader of the Editorial Excellence Team (since January 2022) at Springer Nature. In these roles, he is in charge of leading the continued development and implementation of Springer Nature's research integrity strategy, leading a 20-person team accountable for upholding the highest standards for research integrity in all of Springer Nature's publications. Chris also serves the publishing sector with a voluntary role at the STM Association, the global trade association for academic and professional publishers, where he chairs a committee of senior publishing industry executives overseeing governance of the STM Integrity Hub. The STM Integrity Hub is an initiative launched early 2022 to equip the publishers (and the scholarly communication community) with data, intelligence, and technology to protect research integrity.

Chris has decades of experience in the intersection of scientific publications, academic research, and the ethics of the scientific peer review process. He was a long-time volunteer and is a continuing advocate for the Committee on Publication Ethics ("COPE"), serving in a number of positions from 2005 to 2020 including treasurer, vice-chair, and co-chair, and has worked in a number of other positions at the intersection of research integrity and publishing. He has served on the World Conference of Research Integrity programme committee since 2017. He has written widely on the topic of publishing and academic integrity in the peer review process, including published articles on minimum report standards for life scientists and on how different models of peer review impact final research quality.

Chris started at Wiley in 2004, serving as a Publisher and then an Associate Editorial Director in clinical medicine. He published Wiley's first edition best practice guidelines on publication ethics. In 2011, he was promoted to serve on secondment as the Editorial Director for Health and Life Sciences in Australia. On his return to the UK in 2014 he worked in Wiley's innovation team as the Digital Learning Director and Learned Society Partnership Director, from 2014-2017. His time at Wiley continued with him becoming the Director of Research Integrity in the Wiley Open Research team, a post he held from 2017 to 2021. Chris began his career as an Editor for various publications and holds a bachelor's degree in Chemistry.

Chairman FOSTER. Thank you. And next is Dr. Stell.

TESTIMONY OF DR. BRANDON STELL, NEUROSCIENTIST, FRENCH NATIONAL CENTRE FOR SCIENTIFIC RESEARCH; PRESIDENT AND CO-FOUNDER, THE PUBPEER FOUNDATION

Dr. STELL. Chairman Foster, Ranking Member Obernolte, and all distinguished Members of the Committee, it is an honor to join the hearing today. I'm a U.S. citizen and neuroscientist with the French National Center for Scientific Research and President and Co-Founder of the PubPeer Foundation, which is a nonprofit organization that maintains the website PubPeer.com. My testimony today is my own and does not necessarily reflect the views of the CNRS.

In the fall of 2012, I launched PubPeer with the help of two colleagues, aiming to provide a forum for scientists to discuss the scientific literature. Today, I run it with the help of Boris Barbour, another CNRS researcher. My motivation for the website is to capture discussions of the scientific literature that we scientists typically have in the lab and share them publicly to help other scientists evaluate the scientific literature. The site has witnessed the emergence of a community of expert reviewers and helped expose low-quality research. The numbers—numerous regular PubPeer users are now clearly more experts than most journal and institution staff when it comes to the forensic examination of the literature. We currently receive around 3,500 comments and 700,000 pageviews per month.

Today, I will share insights from the first 10 years of running this website. When we launched this site, we received a flood of comments pointing out serious flaws in the literature. It was as if there was a backlog of unprocessed problems, and the website provided a new release valve to share them. Prior to the website, the procedure for sharing such problems was to write to the authors, their institutions, and the journals and hope that one of them would correct the record. Even when this process succeeded, it was extremely slow, and conflicts of interest often discouraged any action.

The site allows this blockage to be circumvented. Issues can be immediately displayed online so that anyone interested can be made aware and authors can respond. Shortening by months or often years the time it takes to find out about these issues in the literature saves researchers' time and ultimately tax dollars that would have been spent trying to build on flawed research.

How can this be? How did these issues find their way into the literature? Perhaps the underlying reason is not too surprising. With the expansion of science, we continue to create new and important advances at a faster rate than ever, recent examples being development of vaccines during the pandemic and the James Webb Telescope. However, this expansion creates challenges for identifying and supporting the best scientist and scientists—science and scientists.

Job postings for faculty researchers and funding opportunities can now receive hundreds or thousands of applications, and shortcuts for screening those applications become more and more tempting. It's much faster to look up metrics about a journal where an

article is published than it is to actually read the article. And applicants, perhaps falsely, believe that those metrics are the key to the advancement of their careers. In that atmosphere, it's easy to see how many of these issues we see raised on PubPeer find their way into the literature. Instead of publishing a boring result in a journal with a lower metric, there are incentives to select, misrepresent, or even falsify data in an attempt to give a falsely positive result that would land in a journal with higher metrics.

A sensational example of the problem is paper mills, which produce articles for the sole purpose of artificially inflating the publication and citation metrics, while hoping that nobody ever reads them to see that they are fake. They do get published, sometimes by reputable publishers, but perhaps that's not too surprising since journals collect fees for every article they publish, regardless of its quality.

Although I believe these paper mill articles cause little harm to the overall progress of science since they would rarely be confused for real scientific results by scientists, they do highlight the underlying problem with scientific publishing. Incentives need to shift to place higher importance on the content of articles and not on the metrics.

How can we fix this problem? The metrics surrounding articles are now ingrained in the community and unlikely to disappear anytime soon, even if they should. Commentary on sites like PubPeer can provide parallel sources of information that can be much more informative. If scientific commentary continues to grow and involve diverse sections of the community, evaluation committees could start relying on it more than the current metrics so that incentives might shift back toward solid reproducible results that stand up to public scrutiny.

If contributions to this body of evaluation were rewarded when evaluating researchers for funding, promotion, and prizes, it is likely that scientists would participate to a greater extent. The Federal Government through its funding of science could play a huge role, but that potential influence is largely unrealized today. To our knowledge, funding agencies like the NIH (National Institutes of Health) and NSF have no procedures to exploit information available through community curation sites like PubPeer.

In addition to providing evaluation of the publications referenced in grant applications, the information from these sites could be used to reward scientists that make exceptional contributions to public evaluation.

I appreciate the Committee's interest in this very important issue, and I thank you for your time and look forward to answering any questions you may have.

[The prepared statement of Dr. Stell follows:]

Written testimony of

Brandon Stell PhD

CNRS (French National Center for Scientific Research)

President and Co-Founder of The PubPeer Foundation

Before the House Committee on Science, Space, and Technology

For the hearing titled

PAPER MILLS AND RESEARCH MISCONDUCT: FACING THE CHALLENGES OF SCIENTIFIC PUBLISHING

July 20th, 2022

Introduction: History and motivation of PubPeer

Chairman Foster and all distinguished members of the committee, thank you for this opportunity to join the hearing today. I am a US citizen currently living in France and a neuroscientist with the French National Center for Scientific Research (the CNRS), I am also the President and cofounder of The PubPeer Foundation, a nonprofit organization that maintains the website PubPeer.com. My testimony today is my own and does not necessarily reflect the views of the CNRS.

In the Fall of 2012, I launched PubPeer with the help of two colleagues aiming to provide a forum for scientists to discuss the scientific literature, and today I run it with the help of Boris Barbour, another CNRS researcher.

My motivation for the website is to capture discussions of the scientific literature that we scientists typically have in and around the lab, and share them publicly to help other scientists evaluate the literature. I believe that a body of collective public evaluation like this one could also help refocus the assessment of science onto results and data in individual published articles and away from the use of names of journals and secondary metrics that are rampantly used today as proxies for quality of science and scientists. The essential characteristics of PubPeer are: the ability to discuss any scientific article, speed, simplicity, a permanent right of reply for authors, the availability of anonymity and exhaustive content moderation. In the ten years since we launched, the site has grown immensely, thanks to the input

of our users, who have made the site what it is today. The site has witnessed the emergence of a community of dedicated and expert reviewers while exposing low-quality research.

A few statistics help give some idea of PubPeer's scale. We are a worldwide site in terms of the origin of the publications discussed. Our database contains over 130k comments about 45k publications in 4k journals. Although the majority are focused in biology and medicine, many other subjects are represented, such as chemistry, psychology and engineering. Currently we receive around 3500 comments and over 700k page views per month.

Today I will share with you some thoughts and observations from the first ten years of running this website.

Fundamental problems in some of the published literature

When we launched the site I naively thought that comments would be similar to scientific discussions but the overwhelming majority were about more fundamental problems with data. An astonishing number of comments began flooding in that pointed to serious flaws in articles. It was as if there was a backlog of these fundamental problems and the website provided a new release valve to share them without passing through the usual mechanisms for correcting the record. Prior to the website, the process for sharing such problems had largely been to write to the authors, their institutions, and the journals in hope that one of them would recognize the problem and correct the record. However, even when this process was successful, it was extremely slow, and conflicts of interest often discouraged any action from ever being taken, because authors, journals and institutions all have reasons for not wanting to acknowledge or publicize such problems. The site allows this blockage to be circumvented: issues can be immediately displayed online so that anyone interested can be made aware and authors can respond. This process has the advantage of immediately informing those interested in building on a publication (e.g. scientists, policy makers, pharmaceutical companies, etc) about potential flaws, and shortening, by months or often years, the time it takes to find out about such issues saves researchers' time and ultimately tax dollars that would have been spent trying to build on flawed research. Over the last ten years we have received comments pointing out these fundamental flaws in articles ranging from apparently high-profile, ground-breaking research from the best institutions, to articles that were clearly only meant to be fillers for evaluation committees interested in publication metrics. It has become crystal clear to us that the reputation of a journal does not guarantee quality of the research. **It should be stressed that journal and institutional responses remain, on average, very ineffective and inefficient, even when evident problems have been made public, but the site always serves to alert researchers directly, bypassing official inaction.** Furthermore, numerous regular PubPeer users are now clearly more expert than journal and institution staff when it comes to the forensic examination of the literature.

The underlying cause

Perhaps the underlying reason for these issues is not too surprising. With the expansion of science we continue to create new and important advances at a faster rate than ever; recent examples being the ability to rapidly create new vaccines to treat a global pandemic and the James Webb telescope sending back mind-boggling data about the outer reaches of the universe and the beginning of time. However this expansion creates challenges for identifying and supporting the best science and scientists. Job postings for faculty researchers and funding opportunities can now receive hundreds or thousands of applications and shortcuts for screening those applications become more and more tempting. It's much faster to look up metrics about the journal where an article is published than it is to read the article, and applicants (perhaps falsely) believe that those metrics are the key to the advancement of their careers. In that atmosphere it's easy to see how many of the issues we see raised on PubPeer find their way into the literature. Instead of publishing a boring result in a journal with lower metrics, an author could choose to select only experiments that worked, choose less stringent statistical tests or in the rare extreme case rearrange the data in an attempt to give a falsely positive result that would land in a journal with higher metrics. A sensational example of the problem is paper mills, which produce articles for the sole purpose of artificially inflating publication and citation metrics, sometimes by a lot, while hoping that nobody ever reads them to see that they are fake. They do get published, sometimes by reputable publishers, but perhaps that isn't too surprising, since journals collect fees for every article they publish, regardless of its quality. Although I believe these paper mill articles cause little harm to the overall progress of science—in my field they would only rarely be confused for real scientific results—they do highlight the underlying problem with scientific publishing: **incentives need to shift to place a higher importance on the content of articles**.

How can we fix the problem?

If the underlying problem is that the content of scientific articles is undervalued, what can we do to provoke change? The metrics surrounding articles are now ingrained in the community and are unlikely to disappear anytime soon, even if they should. But commentary on sites like PubPeer can provide parallel sources of information that can be much more informative. **If scientific commentary continues to grow and involve diverse sections of the community, evaluation committees could start relying on it more than the current metrics, so that incentives might shift back towards solid, reproducible results that stand up to public scrutiny.** If contributing to this public body of evaluation were rewarded when evaluating researchers for funding, promotion and prizes, it is likely that scientists would participate to a greater extent.

Currently we rely on the journals for curation of scientific literature but there have been many recent encouraging developments from various groups that provide platforms for community curation of the scientific literature in what are known as overlay journals. These platforms enable the scientific community to take control of curation and provide a more democratic system of deciding what are the most important discoveries. The development and growth of such platforms should be encouraged.

Finally, our experience with PubPeer suggests another key intervention. The great majority of disputes on our site would be instantly resolved by access to the original data, which is often unavailable because many publications only show examples and summaries that readers must accept on trust. If data sharing were mandatory, it would strongly discourage low-quality work and misconduct, because data falsification would be harder and riskier, while detection of all problems would be facilitated. But those benefits only accrue if data sharing is required of researchers who wouldn't wish to do it; in other words it should be mandatory.

For all the above, the federal government, through its funding of science, could play a huge role, but that potential influence is largely unrealized today. To our knowledge, funding agencies like the NIH and NSF have no procedures to exploit information available through community curation sites, including PubPeer. In addition to providing additional evaluation of publications referenced in grant applications, the information from these sites could be used to reward scientists who make exceptional contributions to public evaluation.

Legal context

Online forums for discussing science rely on commentary that is often critical and it is not uncommon for scientists to react to this criticism with legal threats to the platform. We regularly receive legal threats[1] from disgruntled authors who prefer not to respond with their data on the website, but because the United States provides us with protection under Section 230 of the Common Decency Act, they rarely become more than threats. Without the protection provided by Section 230 the PubPeer Foundation would be forced to defend these frivolous threats and the legal fees could force us to shut down the website. The SPEECH act also provides partial protection from foreign actions.

1. https://www.aclu.org/cases/sarkar-v-doe-pubpeer-subpoena-challenge

Biography

Brandon Stell is a US citizen who currently resides in Paris, France where he leads a research team with the CNRS (Centre Nationale de la Recherche Scientifique). His team studies a part of the brain called the cerebellum and how it processes sensory information encountered during movement to adapt behaviors to a changing environment. In 2012 he co-founded the website PubPeer.com to provide scientists with a forum to discuss the published research literature. He now runs the website with another CNRS researcher, Boris Barbour, and it has grown to be one of the leading sites for scientific discussion with a dedicated community of users who have helped strengthen the scientific record by exposing and correcting its weaknesses.

Chairman FOSTER. Thank you. And at this point, we will begin our first round of questions, and the Chair will now recognize himself for five minutes.

So first, to all our—all of our witnesses here, in your testimony, you all discuss the problem of incentives. The scientific community, you know, still largely follows the model of publish or perish where simply getting papers out the door is a big component of advancing a career. And so, briefly, what would be involved in generating better matrix—metrics that the scientific community could use to assess the quality of the research? For example, do the metrics adequately, you know, punish the ratings of researchers that have a high rate of retracted papers? Or maybe the traditional H index, you know, which I presume has been badly corrupted by paper mills you know, maybe it should—instead of being recorded as a positive integer should be recorded as a complex number with both a real and imaginary component for the H index.

So if you could just speak briefly on what improved metrics are under under consideration at this point, and we'll go in reverse order. And I'll start with Dr. Stell.

Dr. STELL. So I personally feel that metrics in general are a horrible way of evaluating science. I think that we need to, as a community, put more focus on the content of the articles, read them, and evaluate scientists for what they publish and not the metrics associated with those articles. And I think that one way of accomplishing this is by creating this body of information from scientists that are commenting on articles and creating a secondary source of information in addition to these metrics that will hopefully overtake the use of metrics one day.

Chairman FOSTER. And, Mr. Graf?

Mr. GRAF. Thank you. Yes, that's a good question. The things that I read about research assessment will point in the same sort of direction that Brandon just described, toward a more qualitative approach to understanding the impact of a piece of research. And in some nations that—I believe that's done through descriptive case studies where researchers prepare a report about a piece of work that explains its importance. And then the Research Assessment body uses that, perhaps with a suite of metrics, not relying on one, but with multiple metrics to form an opinion and to reward that researcher. So I definitely think the mood in the room is a distinct desire to move away from reductive metrics and toward something that's a lot more reflective, the complex nature of research itself, and much more qualitative.

Chairman FOSTER. Yes, Dr. Byrne?

Dr. BYRNE. Thank you. Just to add to those comments, I think it's helpful if there are publication requirements leveled on professionals, that those requirements or expectations recognize people's training and capacity to conduct research. Clearly systems where, for example, medical doctors are required to publish papers when they have neither the time, the resources, or the training to conduct research, that seems like a bad system. I would prefer a medical doctor to be assessed based upon their capacity to care for patients and provide cures.

In terms of other forms of metrics, I think a number of commentators have thought about the desirability of the journal impact

factor to also reflect other dimensions of publisher activity such as the capacity to correct the literature, as well as to produce it. Thank you.

Chairman FOSTER. And so another thing that I think has occurred to probably all of us, reading your testimony, is the role of government in supporting more and better automated tools on this. Is that—you know, I loved, Dr. Byrne, your analogy with waste removal, that we just need a system for waste removal in the system. And so are—and we—you know, in normal communities throughout the country, part of the budget goes into waste removal, and maybe part of the solution is just turning up the fraction of our municipal budgets we would devote to that.

Are there specific proposals that you've seen that seem to make sense for—you know, for putting more muscle into this effort?

Dr. BYRNE. Look, I think—I've thought about this a lot, and it is unusual that science is so geared toward publication or production and sort of relentless production without that capacity to just cleanup every so often. So there could be a system where, you know, a certain proportion of research budgets could be devoted to that particular activity. Another system could be that if you are given research funding to do original research, that perhaps a proportion of that budget could be also devoted to a certain kind of quality improvement activity within the literature.

Chairman FOSTER. Thank you, and my time is expired. But I love the idea of financial incentive somehow. I've learned and—when I moved from science to politics is that people respond awesomely to them, so that if there was some bounty on taking down garbage papers, I suspect we've got a very active international community collecting that bounty.

And so at this point, I'd like to recognize the Ranking Member for five minutes of questions.

Mr. OBERNOLTE. Thank you, Mr. Chairman. And thanks to the witnesses, a really interesting topic here.

So, you know, I think it's always interesting to concentrate on the—you know, the most negative consequences of the problem that we're discussing. So the attorneys call that the parade of horribles, the worst thing that could possibly happen. And in my opinion, when you're talking about paper mills, we've got two potential consequences. You've got fraudulent research getting rewarded by unqualified people graduating or by unqualified people earning promotion, right? So we can attack that perhaps with tools to identify these fraudulent papers and hopefully create consequences for the people that are using the paper mills.

But to me, the more consequential problem is the impact on research integrity, this idea—and I think Dr. Byrne brought it up a little bit—that we're—researchers might be misled in their daily work by the presence of these fraudulent papers in the body of literature.

But I want to tunnel down a little bit on the likelihood that that would happen. And I'd like to open this up to anyone that wants to comment of our witnesses on this because my first question is, wouldn't—if you were an editor of a journal and you're reviewing a paper that was submitted and it came from a paper mill, wouldn't it be pretty immediately apparent? I mean, right now, the

paper that Dr. Foster generated, you know, that—anyone with a technical background could read it and realize that there were inconsistencies. But, you know, we're looking at maybe one more generation ahead, you know, where AI tools might create something that would convince someone that wasn't an expert in the field. But if you're an editor in a journal, I mean, you're familiar with what the state-of-the-art is in various fields. Wouldn't you immediately know, you know, reading a paper, well, this doesn't make any sense because, you know, really what people—the topics of research are A, B, and C, and, you know, this isn't even consistent with that. Wouldn't people know about that?

Dr. BYRNE. Thank you. Look, I might just—I'll stop and just give a brief answer. I think that some paper mill papers are actually highly plausible. They have already been accepted for publication, which means they've passed editorial review, they've passed peer review, and they've moved into the literature. So they're actually highly plausible, and they very closely resemble genuine papers. And I think that's the great danger. I mean, I can certainly think of many researchers who would read the kinds of papers that we study, particularly students and early career researchers that have not spent 30 years reading the literature as I have. These papers are highly plausible, and they are capable of misleading people purely because of that.

Mr. OBERNOLTE. Interesting.

Dr. BYRNE. Thank you.

Mr. OBERNOLTE. So a follow-on question, you mentioned peer review. And I am astonished that peer review wouldn't help us take care of this problem because, ostensibly, if you've got a panel that's doing peer review on a paper, those people are familiar with, you know, the cutting—the—where the cutting edge of the research in that field is. I mean, I remember when I was in graduate school. For, you know, one brief shining moment, you're supposed to be the world's expert at a certain very narrow field, and so one brief shining moment, you know, I knew everything there was about public sector budgeting and research on various budgeting methodologies because that's what my dissertation was on. And at that moment if you had given me a paper on something related to that field, I would have been able to say, wait, I know there hasn't been any research on that at all and this doesn't actually make any sense because here are the topics that people are researching. So doesn't good—a good peer-review system help us fix that problem?

Mr. GRAF. May I take up the response there?

Mr. OBERNOLTE. Dr. Graf, go ahead.

Mr. GRAF. If I could try and tell a brief story to explain why the—how the paper mills navigate around both editors and peer reviewers, that might help. So, last year, I became aware of an editor-in-chief of a journal who'd become concerned about the volume of content being published and the scope of the content being published in a guest-edited issue that he had appointed a guest editor for. He emailed the guest-edited issue, who he thought was—guest editor who he thought was the guest editor for that issue. And the guest editor, she responded saying this has got nothing to do with me. So then we checked the email addresses and noticed that the guest editor was using an email address very close to the email ad-

dress of the legitimate researcher who the editor-in-chief thought he'd appointed but not the email address. And so our assumption is probably the truth, that a paper mill has—was using identity theft and identity fraud to place a fake guest editor in charge of this guest-edited issue. That fake guest editor was then appointing fake peer reviewers to fake peer review the content and passing the content through as if it were legitimate and essentially, you know, then triggering the publication of that content.

So paper mills are devious and have worked ways around the largely trust-based and professional courtesy-based system that has been working quite well for decades, if not hundreds of years, and we need to tool up to prevent that better.

Mr. OBERNOLTE. Right. Well, it's interesting. You just raised a—well, perhaps we'll have a second round here. I see I'm out of time, and I don't want to abuse the process. But it's interesting that, Mr. Graf, you've just raised, you know, kind of another follow-on problem, which is the corruption of the peer-review process, in the service of these paper mills, which, you know, I think is—you know, I put that up, along with the other consequences of what we're discussing here. But thank you very much. I'll yield back, Mr. Chairman.

Chairman FOSTER. Thank you. And I think we will attempt to have a second round of questions if time and Members' attendance allow.

We'll now recognize Representative Casten for five minutes.

Mr. CASTEN. Thank you so much. This is a fascinating conversation. I want to just start—I want to understand from Dr. Byrne in your research, do you have any sense—and I realized that you've—you're trying to quantify it. Do you have any sense of how much of the issue is plagiarism versus fraud?

Dr. BYRNE. That's a very good question. We mostly—we don't really study plagiarism in detail. My background as a molecular biologist is the kinds of reagents that are used to construct experimental results. So we study the reagents that are used and whether those reagents are correctly identified. Sometimes we do see reagents appearing across model papers when that seems highly unlikely, so that indicates that there may be some role for plagiarism. But we feel that it is more likely that most paper mills are creating papers according to templates, so they have a—kind of a basic skeleton of structure and then they fill in the gaps in different ways. And they often——

Mr. CASTEN. OK. And——

Dr. BYRNE [continuing]. Target topics that are not very well understood and so peer reviewers have no knowledge of these papers and so can't critique them.

Mr. CASTEN. And I ask this because it seems like the the tools that we might have to address them are different, you know, and there's ways to detect fraud, and there's different ways to detect plagiarism. I'm curious for your thoughts—and I guess this would apply in either case. There certainly has been some talk in the—you know, especially in sort of the bioethics community about should we mandate that researchers publish negative results since there's no real incentive to do that? And all of us who have ever worked in a lab knows that most experiments don't get you pub-

lishable results. If—and I don't know how we do that. But if we were to mandate that—you know, that labs or researchers had to publish, you know, all their experimental results even if they were negative, would that help or hinder this problem?

Dr. BYRNE. Look, I don't know. I think that the publication of negative results is very important, but one of the issues around publishing negative results is all of these publications take time, and I think it will be difficult to incentivize that process. I mean, researchers, like everybody, are more interested in something that they are intrinsically interested in, and sometimes they don't find negative results particularly interesting. But I agree that I think we have to find ways of removing this incessant focus upon results that must be positive, and we need to be teaching our students that a negative result is just as important as a positive result. It's a result.

Mr. CASTEN. Yes, I mean, I wonder sort of, you know, to use a bad baseball analogy, if I only knew that a hitter—if I only knew the stats on a hitter when they got hits, I might not—you know, I wouldn't know the difference between Ted Williams and, you know, whoever's the third string on a baseball team. And so if I knew a researcher was batting 800, I'd be a little skeptical, right? Anyway——

Dr. BYRNE. Yes, that's——

Mr. CASTEN. The—this is a very wonky one, and I suppose this would only work for falsified data. There's this wonderful little numerical trick that, you know, in any sequence of numbers that sequential, if you look at the probability of the digits, ones are more likely than twos, twos are more likely than threes. Is there a way to automate that? You know, because I would imagine if I'm—you know, if I'm reporting the number of colonies in an agar plate, that's a very hard thing to fake and—but it's algorithmically testable. Is that worth time? Are people already doing that? Is that just a dumb thing that I read about years ago that's not relevant today?

Dr. BYRNE. No, no, it's certainly not a dumb thing. I mean, people are certainly looking at that kind of thing, particularly clinical trial data, patient data, where it's actually very, very difficult to fake random data. So I think the answer is that we need different kinds of tools for different kinds of data and different kinds of science. And we don't have all of those tools. We have some of them now, but we don't have them all.

Mr. CASTEN. So my last question is sort of deeply philosophical, and I'll try to get this off in a minute. It's always struck me that whether you're doing basic science, you know, at the lab bench or doing research, you know, going through the literature review, I always thought Immanuel Kant got it right, that, you know, all you can do is prove things wrong, you can never totally prove them right. And things are more likely to be true when you try to falsify them and fail. And I think our human brains are really good at that kind of analysis, you know, the—our ability to say, well, if this thing is true and the causality arrow points in this direction, then that would imply that this other thing is true. Let me see if that's the case. And if not, I got some problems over here. I don't know how you write algorithms to do that. I think our brains are just

sort of uniquely set up to do that. And it's what the traditional peer review process is really good at.

Is that even a—is it even algorithmically possible to do that sort of Kantian falsifiability? And if so, or if not, is there any way to sort of satisfy that a paper that is deemed worthy of publication someone has attempted to do that falsifiability and failed? Does that make sense? I realized I'm getting very philosophy of science there, but does that make sense as an approach?

Dr. BYRNE. Look, it makes sense. I don't know if it's possible, but that's—I'm not an algorithm person, so, you know, perhaps one of the other witnesses could answer.

Dr. STELL. So yes, I think that this would be possible. This sort of thing might be possible. But I think that the more algorithms we build, people are going to find ways around these algorithms. And so I think that, you know, this is maybe not—shouldn't be our focus is trying to find every individual instance of fraud but changing the incentives so that fraud is no longer a winning strategy, that we put the focus on the content of the articles and make it so that it's just not important to do fraud anymore.

Mr. CASTEN. Thanks. And I yield back.

Chairman FOSTER. Thank you. I'll now recognize Representative Bice for five minutes.

Mrs. BICE. Thank you, Chairman Foster, and thank you for the witnesses for being with us this morning.

My first question really is to any of the witnesses, and that is how long do you think that paper mills have been impacting academia? And how many of them do you believe that are based in Russia or China?

Dr. BYRNE. I can start. In terms of how long they've been operating, I don't think we have a clear answer for that. I would estimate that at least since 2008, so possibly for about 15 years. In terms of the numbers of paper mills, again, if I refer to the literature, a paper was written in 2013 that estimated the number of of ghostwriters that were operating in China at that time in 2011, they estimated nearly 1,800 full-time-equivalent ghostwriters in China. There's been very little research done on this topic, and so I think the answer is—you know, the answers today are not clear.

Mrs. BICE. Thank you. Mr. Graf, did you want to follow up with that?

Mr. GRAF. Thank you. I could add a little. I think it's true that the strategies and tactics that paper mills have been using since data Jenny Byrne cited have changed. In my—the information I've studied, the use of AI and algorithms to generate text in articles probably began around 2019, and that's when you've got a massive change really in the tools that paper mills have got available to them. So that talks both to sort of how long and also to change in their practices, which refers back to what Brandon said, which is the sort of arms race thing. It really would be—it would be beneficial to get out of the arms race because—and really address the incentive system. I think that's probably the right way forward.

Mrs. BICE. How do you—to follow up on that particular point, Dr. Stell, how do you—what do you think we should be doing to try to address this in a way that still allows for researchers to be able

to publish papers, but it's not solely focused on the data? Where's the fine line there to be able to publish a research paper that has value, that can be utilized by the community, but yet isn't solely—isn't isn't focused on specifically the data that's contained within it?

Dr. STELL. Thanks for the question. Yes, I think that anything we can do to put more focus on the content of articles is going to help enormously. And I think the one thing that hasn't been done that could be done is rewarding people for contributing to a body of commentary. If we have this body of commentary, this is going to take the focus off of metrics and put it on expert opinion of scientists. And if we can reward expert scientists for their commentary, we're going to get more participation, and then these evaluation committees are going to start using that expert evaluation. And so I think that for me the way forward is to just create another body of evaluation parallel to metrics, which are going to continue to exist for a little while but perhaps could be replaced by more expert informative opinion.

Mrs. BICE. And on Dr. Graf's—I'm sorry, Mr. Graf's point and, Dr. Stell, you're welcome to comment as on this, but I think that the—one of the concerns I have being a Member of the House Armed Services Committee is how do we ensure that these research papers aren't being influenced by foreign governments in a way that could have a negative impact on security? Any thoughts there?

Dr. STELL. Go ahead, Chris.

Mr. GRAF. No, I think it's a question about misinformation, isn't it? And I don't think that the motives behind all of this are to promote misinformation. I think they're the simplest of motives. Paper mills want to earn money. They earn it when researchers give it to them. Researchers want to earn money. They earn it when they get a paper published and when they get promoted. So I don't know. I don't have evidence to suggest there's a conspiracy from foreign agencies going on.

Mrs. BICE. You don't think that there's sort of falsification in a way that American academia would be impacted by that falsification?

Mr. GRAF. I don't think that the motives are to impact American academia, no. I think the motives are very isolated to the individual researcher who's buying this bogus service from this—from the paper mill.

Dr. STELL. Can I jump in?

Mrs. BICE. I think my time has expired. I yield back.

Chairman FOSTER. Well, thank you. And I think we will now start a second round of questions, and so I will recognize myself for five minutes.

So I'd like to return to the subject of trying to get the financial incentives right. Because, as Mr. Graf pointed out, that really drives a lot of this. You know, I've often fantasized that if it cost people 50 cents to send me an email and then they got the 50 cents back if I saw fit to respond to it, that my spam filter would have a lot less work to do. And so, first off, has anyone estimated how many man-hours would be actually required to adequately peer review this flood of papers and what the total cost of that would be and the cost per paper?

Mr. GRAF. Not to my knowledge.

Chairman FOSTER. Even order of magnitude? OK. If you can just reply for the record, I'd be interested in that. Because that's—you know, it seems like a daunting number of man-hours, and you want to engage the best and brightest in any field toward reviewing those papers, and it's a huge burden to put on their time and might not be the best use of their time.

Mr. GRAF. I can add a little information, but it's not really about peer review. It's about the internal review in an investigation of potential paper mill papers that I've been conducting with my team. And that review has been ongoing since September last year and has taken—I hadn't added up the person-hours, but there's been a team of, let's say, five people. It's more than five, so maybe it's between five and 10 people working at least part-time, sometimes full-time, on the exercise, and they have spent money on consultants as well. So—and that's been focusing on a total kind of universe of about—where we looked at about 3,000 papers, so that's not looking at the whole world, right? That's only looking at a small part of the world. And there you go. That's the sort of sketch that I've got for you, the amount of—for you about the amount of effort that it's taken post-publication, not with the peer-review community but with colleagues internally at a publishing company.

Chairman FOSTER. Now, would a more significant cost to submit a paper for review be an effective partial remedy, or would that really place an unacceptable burden on emerging researchers? You know, for example, if—you know, when you got a grant, five percent of the grant money you could allocate toward getting any publications reviewed, and then you'd spend that money where you thought it would do the most good. Are there incentives like that that could be put into place?

Mr. GRAF. That's an interesting idea.

Chairman FOSTER. Has that ever been talked about?

Mr. GRAF. Publicly, there's a couple of different campaigns to—led—there's one led by James Heathers to—a gentleman called James Heathers to claim payment for peer review. And perhaps if—I don't know how much sort of ground that movement has been made—has made. I do worry about equity and access to publishing services.

Chairman FOSTER. Yes, so presumably, this would be based—you know, the fees will be based on ability to pay based on your situation and your country's situation.

Mr. GRAF. Yes.

Chairman FOSTER. But—and that—but maybe that's a partial solution. You also mentioned identity fraud, and there is a lot of progress. In fact, where I'm marking up this week a bill that we're pushing forward to secure digital identity. This is just providing tools for individuals to prove they are who they say they are online and also to attach verifiable credentials to that digital identity so you can't fraudulently claim, you know, basically, credentials that you haven't earned. And so is that—is there—are there things underway in the academic community already? It also provides a mechanism for punishing people that abuse the system or at least identifying them so that they can be dealt with with appropriate suspicion? Are there any—anything along those lines underway?

Dr. BYRNE. I think I can just speak to that briefly. So there's a system of author identity called ORCID (Open Researcher and Contributor ID) that has been running for some time. Some journals are now requesting that all authors have ORCID identities as a way of combating paper mills. But I think paper mills are very adept at getting around these kinds of fairly small hurdles that we place in front of them. So there is evidence that paper mills then simply take out ORCID identifiers for their potentially real or fake authors. So that's a major issue, I think. I think that probably also pertains to paying fees for submitting manuscripts. Paper mills would probably be willing to pay those fees.

Chairman FOSTER. OK. But presumably when you claimed an academic credential from—and then attempted to attach it to a fake identity, at some point, the university whose academic credentials are being stolen would blow the whistle on you.

Anyway, I will—I'm out of time here, and I'll yield to the Ranking Member for five minutes.

Mr. OBERNOLTE. Well, thank you, Chairman Foster.

I want to talk a little bit more about research integrity because I think that that is the most dangerous consequence of these false and fraudulent papers floating around. So, you know, the presence of misinformation is an issue that we are dealing with as a larger society. It's not just academia, and it's not just research, certainly no social media. There's lots of information out there that's true, there's lots that's not true, and there aren't a lot of tools for a user to try and figure out what's—how to differentiate between those two.

But we've got a tool in academic literature that is not available to social media, and that is the hierarchy of different scientific publications. I mean to call something literature ignores the complexity of the fact that scientific publishing is not monolithic. And we've got journals that are highly respected in their field and then journals that are less so.

So, you know, a question for anyone who would like to answer. Does—is there a way we can leverage that hierarchy? You know, the fact that there are journals that can serve as highly trusted sources of information, can we leverage that to help us solve this problem?

Mr. GRAF. If I may start, one of the things that we can do is really get behind the transition, or transformation even, away from subscription publishing to open-access publishing. When we make all of the more trustworthy information that's available in journals, including those of the top end of the hierarchy, more open, it's there as a counterbalance to the misinformation that is freely and openly available on the web. So I think there's an argument there for that transformation to open.

Mr. OBERNOLTE. Well, Mr. Graf, I agree that that would be great, but, I mean, there are monetization problems with doing that. How would you solve those? How would you alter the monetary incentives that empower this current subscription model?

Mr. GRAF. That's a whole query of its own. But yes, we're working on that as—across the research publishing sector, and we're, you know, intent on making the transformation—moving the money that is currently being spent on subscriptions into a way to

enable the—those journals to then be open. So it's—yes, it's complicated. And one type of openness won't suit all research disciplines and all journals or regions on the planet, but that's our goal.

Dr. STELL. One thing I would just like to add to that is that there have been studies looking at the impact factor of a journal and the number of errors that are found in the journal. And the higher-ranking journals are not necessarily immune to the problem. So I think if we start relying on these tiers to tell us which is reliable information, it's not as accurate as we would hope it would be.

Mr. OBERNOLTE. Interesting. I wouldn't have guessed.

Dr. Stell, while I've got you, one of the things that I was fascinated by is in the previous round of questioning you were talking about how some papers, fraudulent papers had been identified on PubPeer. And I'm wondering, you know, as we grapple with this issue of eliminating the commercial and academic incentives for publishing fraudulent papers, what were the consequences for the authors of the papers that your website identified? Did they have consequences?

Dr. STELL. There are examples of absolutely zero consequences, but there are also examples of people being fired from their positions for having been caught cheating and exposed on PubPeer. So the consequences range.

Mr. OBERNOLTE. So when people—when there were consequences, what triggered the consequence? So if someone on PubPeer looks at a paper and says this was generated by a paper mill, here's my proof, here are the, you know, inconsistencies. And so, you know, the PubPeer community agrees this paper is fraudulent, does someone reach out to the employer of the person, the author of this fraudulent paper and say that, you know, if they cited this paper in their resume, then they were hired under fraudulent, you know, circumstances? I mean, did someone take that affirmative action?

Dr. STELL. That's a very good question. We're not part of that process. What we do is we provide a platform for people to discuss these things and make it public. And I have to say that paper mills are a real minority of the discussions that happen on PubPeer. But the fraudulent work is usually some sort of image that has been copied, some data that has been misrepresented. It's been exposed on PubPeer. That is public for everyone to view, including the people's employers and other committees. And so it's—presumably, it's taken up by those people, and they're taking actions. It's not—we're not part of anything other than making that information public.

Mr. OBERNOLTE. Right. Well, I see I'm out of time. I want to thank everyone for the, you know, really interesting and consequential hearing here. But let me reiterate my conviction that when we're talking about the spread of disinformation, and academic literature is no exception, I think that trying to focus on eliminating all fraudulent papers is a fool's errand because I think they're going to be out there, so I think the better solution is to try and create trusted sources of information where some peer-reviewed vetting takes place and people can have a higher degree of trust in. And, you know, to your point, it's clear that more work

needs to be done there. But, you know, certainly I think that's where the solutions are going to be found.

Anyway, thank you, everyone, for the discussion. It's been really interesting. I yield back.

Chairman FOSTER. Thank you. And I just want to second the Ranking Member's comments about the importance of trusted sources. You know, it used to be—I—at least, I felt when I was in my career in science, if you stood up and said something that you know was not true, that that was a career-ending thing. And now, you know, as I moved from science to politics, it seems like you stand up and say something that's not true, it seems to only increase your chances of reelection. And so we have to do a better job, and science should continue to lead by example of taking a very hard line when outright fraud is detected. And I was encouraged to hear that at least some universities are—you know, you're losing tenure and you're out of here if you participate in any of this because that—it's important. Science always operates at the edge of what is known, and we cannot tolerate deliberate lying when we're trying to flesh out the details of nature's complexity.

Now, before we bring this hearing to a close, I wanted to thank again our witnesses for testifying. And the record will remain open for two weeks for additional statements from the Members and for any additional questions the Committee may ask of our witnesses.

So the witnesses are now excused, and the hearing is now adjourned.

[Whereupon, at 11:07 a.m., the Subcommittee was adjourned.]

Appendix I

Answers to Post-Hearing Questions

ANSWERS TO POST-HEARING QUESTIONS

Responses by Dr. Jennifer Byrne

1

Additions to the written testimony of
Jennifer Anne Byrne PhD
Conjoint Professor of Molecular Oncology
Faculty of Medicine and Health
University of Sydney, Australia
Before the House Committee on Science, Space, and Technology

For the hearing titled:

"Paper Mills and Research Misconduct: Facing the Challenges of Scientific Publishing"

July 20, 2022

Responses to Questions for the Record

Question from Chair Bill Foster during the hearing on July 20, 2022, taken on notice

(From the transcript, verbatim): And so, first off, has anyone estimated how many man hours would be actually required to adequately peer review this flood of papers and what the total cost of that would be and the cost per paper? Even order of magnitude? Okay. If you can just reply for the record, I'd be interested in that.

This is a very pertinent question. The cost of peer review was recently examined by a team of researchers from Hungary and Australia, which included Professor Alex Holcombe from the University of Sydney, with which I am affiliated. Their results were published in 2021 in the journal *Research Integrity and Peer Review*: "A billion-dollar donation: estimating the cost of researchers' time spent on peer review" (ref. 1).

Globally, it was estimated that in 2020, researchers contributed over 130 million hours of unpaid time as peer reviewers, which equates to almost 15,000 years of constant work (ref. 1). *The US was estimated to be the largest single contributor, where US-based peer reviewers contributed over US$1.5 billion in unpaid time* (ref. 1). These estimates, which were indicated to be conservative, highlight that global unpaid peer reviewer contributions are likely to be worth billions of dollars per year. Aczel and colleagues proposed different ways to reduce these costs (ref. 1). These include increased sharing of peer reviews of manuscripts that are resubmitted to new journals (ref. 1). Paper mills also contribute to waste in peer review. We have suggested increased use of preprint servers to reduce unethical duplicate manuscript submissions from paper mills (ref. 2). This could reduce immediate burdens upon peer reviewers while screening methods improve (ref. 2).

References

(1) Aczel B, Szaszi B, Holcombe AO. A billion-dollar donation: estimating the cost of researchers' time spent on peer review. *Res Integr Peer Rev.* 2021; 6(1):14. doi: 10.1186/s41073-021-00118-2.
(2) Byrne JA, Christopher J. Digital magic, or the dark arts of the 21st century-how can journals and peer reviewers detect manuscripts and publications from paper mills? *FEBS Lett.* 2020; 594(4):583-589. doi: 10.1002/1873-3468.13747.

Questions submitted by Chair Bill Foster

1. In the hearing we discussed the development of automated tools that can help journals and peer reviewers detect research misconduct in scientific papers. You noted that there is little funding devoted to research to improve the detection or correction of fraud in publications.

a. Do you think there is a role for the Federal Government in funding the development of new, improved tools for detecting research misconduct? If so, is this something the National Science Foundation should lead, or a coalition of Inspectors General, or someone else?

Thank-you for this important question. Yes, I believe that there is a role for the Federal government and governments in other countries to support development of new and improved tools for detecting research misconduct. The US Office of Research Integrity (ORI) has funded our research to develop and apply the Seek & Blastn tool (ref. 3) and this support was instrumental in obtaining further grant funding from the Australian government. The ORI also funded the Computational Research Integrity conference that was organised by Dr Daniel Acuna at Syracuse University in March 2021. These opportunities could be expanded to provide dedicated opportunities for tool development and improvement. Dedicated opportunities would improve the profile of this research and provide incentives for new people to enter the field, through for example workforce re-training opportunities (ref. 4). Tool development can also be enabled through competitions that can also encourage the formation of new teams that include students and early career researchers.

References

(3) Labbé C, Grima N, Gautier T, Favier B, Byrne JA. Semi-automated fact-checking of nucleotide sequence reagents in biomedical research publications: The Seek & Blastn tool. *PLoS One.* 2019; 14(3):e0213266. doi: 10.1371/journal.pone.0213266.

(4) Byrne JA, Grima N, Capes-Davis A, Labbé C. The possibility of systematic research fraud targeting under-studied human genes: causes, consequences, and potential solutions. Biomark Insights. 2019; 14:1177271919829162. doi: 10.1177/1177271919829162.

b. Another concern is how to evaluate the performance of various tools for detecting misconduct in scientific publication after they are developed by the private sector, the government, or a university. Does the government have a role in testing or validating automated detection tools? Is there a technical standards organization that might be doing this already or might step into this role in the future?

This is an excellent question that covers several important issues. It is vitally important that tools to identify possible misconduct in scientific publications are rigorously validated by independent expert teams. To date, there appears to have been more researcher interest and/or funds available to build new tools, compared with the support available for tool testing and validation. For example, a recent review of 27 methods to assess research misconduct in health-related research found that no method had been formally validated (ref. 5).

Government could provide funding that incentivises tool testing in an independent, rigorous and transparent manner. Dedicated funding opportunities could support both individual tool validation and inter-tool comparisons. There is an important need to test different tools in combination, as the use of multiple methods is recommended to detect potential research misconduct (ref. 5). These funding opportunities should include developing or repurposing "gold standard" publication collections that serve as standardised testing materials. These publication sets allow iterative versions of the same tool to be compared side by side in controlled experiments. The same publication sets can also be used to compare different tools, so that their relative or combined strengths and weaknesses can be identified and addressed. Once these publication sets have been defined, they can be reused many times.

I am not aware of any technical standards organisation that is already overseeing the testing of tools for detection of possible research misconduct, but I agree that such organisations could represent appropriate oversight bodies.

Reference

(5) Bordewijk EM, Li W, van Eekelen R, Wang R, Showell M, Mol BW, van Wely M. Methods to assess research misconduct in health-related research: A scoping review. *J Clin Epidemiol.* 2021; 136:189-202. doi: 10.1016/j.jclinepi.2021.05.012.

Questions submitted by Representative Suzanne Bonamici

1. A recent article[3] in the journal Science revealed that a seminal paper[4] on a potential cause of Alzheimer's Disease contained fraudulent images. That article is likely to have contributed significantly to the direction of Alzheimer's research over the past 15 years, demonstrating the harm that a single fraudulent paper can do to the research enterprise. I understand your research has included detections of fraudulent papers within the pre-clinical trial literature.

a. How significant were the fraudulent papers you identified, as measured by citations or any other appropriate metric of scientific significance?

Thank-you for this important question. In our 2022 Life Science Alliance paper (ref. 6), we found that 712 problematic human gene research papers that we identified had been cited over 17,000 times, as of March 2021. This was an unexpectedly high number of citations, given that most papers were published in specialty journals. It's important to highlight that we do not yet know for certain which of these 712 human gene research papers could be fraudulent, as some errors almost certainly occurred in the context of genuine research. However, the repetitive features and at times impossible experimental results that were described by many of these 712 papers suggests that some may have been constructed with assistance from paper mills (ref. 6).

These 17,000 citations contradict assumptions that erroneous papers attract only limited attention. As further unpublished research from our team has shown that most citations of problematic papers are positive, this indicates that most researchers accept their results, and do not recognise these papers as questionable.

b. Can you directly link any of the fraudulent papers you identified to a specific clinical trial?
This is another very pertinent question. We indeed found that some problematic human gene research papers had been cited by clinical trials, and that more papers (approximately one quarter) were predicted to be cited by future clinical research (ref. 6). In unpublished results, we have found further examples of problematic pre-clinical research papers that have been cited by clinical trial publications, including clinical trials for cancer and rheumatoid arthritis.

Reference

(6) Park Y, West RA, Pathmendra P, Favier B, Stoeger T, Capes-Davis A, Cabanac G, Labbé C, Byrne JA. Identification of human gene research articles with wrongly identified nucleotide sequences. *Life Sci Alliance.* 2022; 5(4): e202101203. doi: 10.26508/lsa.202101203.

Responses by Mr. Chris Graf

SPRINGER NATURE GROUP

Springer
Science+Business Media, LLC
One New York Plaza, Suite 4600
New York, NY 10004-1562, USA

T +1 212 726 9200
www.springernature.com

Answers to questions for the record (QFR) for the House Committee on Science, Space, and Technology, Subcommittee on Investigations & Oversight, hearing:

Paper Mills and Research Misconduct: Facing the Challenges of Scientific Publishing

Witness: Chris Graf, Chair, Governance Committee, STM Association Integrity Hub, and Research Integrity Director, Springer Nature
Dated: Wednesday August 24, 2022

Contents

Questions for the record 1
- **Question 1a** 1
- **Question 1b** 2
- **Question 1c** 3

Questions for the record

Thank you: Thank you Chairman Foster, Ranking Member Obernolte and the esteemed members of the committee for the opportunity to present on how the publishing sector is responding to the challenges of research misconduct. It was a privilege to have been part of the hearing, and it is a privilege to now respond to the questions for the record, below. I look forward to continuing to serve in any way that might be most useful for you.

Statement that accompanied the sub-committee's questions: Identifying fraudulent papers can be challenging and time-consuming work, requiring both subject matter expertise and general knowledge of fraud.

SPRINGER NATURE GROUP

Question 1a

Question: Are there any estimates of the time or resources required to review the entire literature for fraud? What about reviewing a single journal or another logical subdivision of the literature?

Answer: The pre-publication work publishers have done to check for concerns prior to publication -- including by organizing peer review -- means that published research was judged at the time of publication by peer-researchers and scientist-editors to have been sufficiently robust for publication. Publishers invest billions of dollars and countless hours in this process, in partnership with peer-reviewers who contribute time worth hundreds of millions of dollars each year (see this report or this paper).

Hypothetical retrospective re-review of everything ever published would be a potentially enormous undertaking. This would be not only a technical challenge, but also a human challenge. The variety of dimensions that could be checked after publication for concerns is potentially as many as the types of research that exist. It would be insufficient to rely only on computer-generated evaluation; people would be needed to interpret and validate assertions generated by software, and -- by current standards adopted across the publishing sector -- attempted communications with authors would be required.

The requirements for what makes research sufficiently robust for publication do change over time. Publishers continue to invest in screening for integrity, including routine checks for plagiarism as well as for other indicators of ethics and quality (like disclosures of conflicts of interest, description of ethics committee approval and funding sources). That is evolving. Some publishers are beginning to roll-out screening for newer concerns that indicate paper mills, like image manipulations (apparent image manipulations may also result from a data-management error, or naive efforts to make images look presentable; see 'The fight against fake paper factories,' Nature). Both existing and new screening initiatives need investments in a mixture of technology and the people, more jobs in research publishing, continued reliance on and I'd argue strengthened connections with scientific communities. Publishers can and are doing more to stop publication of papers that present integrity risks. However, the opportunities exploited by paper mills, along with other integrity concerns, are created 'upstream' where research is done. Other actors also have a responsibility, including the organizations that fund and employ and set policy for researchers and research.

Question 1b

Question: During the hearing you mentioned that you were part of a team examining a specific subset of about 3000 papers. Approximately how many person-hours did that project require? Do you believe that number would hold constant across a larger body of papers or carry over to papers in a different field?

Answer: We have accounted for the person-hours required for an investigation of 350 papers that we believe were published by a paper mill. The paper mill had inserted an imposter into a position as a 'guest editor' for a journal special issue. The investigation took 75 hours over about a month to confirm findings for each paper involved. This total excludes further time spent after findings were confirmed to attempt to communicate with authors, evaluate their responses, finalize decisions, and then publish retractions. With respect to extending across a larger body of papers and/or into different disciplines: This investigation was looking at a known cohort of papers for specific concerns; different concerns investigated in other situations would require different forms of analyses, which may take more or less time to perform. I would not use the time we spent to estimate the time required for other types of investigations.

I believe that the part of the hearing where this topic was discussed was referring to the volunteer-efforts of readers to identify concerns in published papers more than it was referring to the kinds of efforts employed by publishers. I have no estimate for the time that readers take doing this. Publishers, myself included, are genuinely grateful for the work of readers who identify concerns in published papers. On receiving notification of a concern from a reader publishers almost always need to validate and verify the concern, and then take steps to communicate with authors (perhaps also their employers, i.e, universities and research institutions), evaluate their responses with the journal editor, finalize and communicate decisions, and then take the appropriate action (which could be to publish a retraction, or could be something else including no action). In my experience readers are sometimes frustrated by the time required to take these necessary steps. For concerns like paper mills, I think new ways of working that match the scale and type of the problem are needed. This is one of the key recommendations from the paper mills research report published by COPE and the STM Association earlier in 2022.

Question 1c

Question: Detecting fraud can require someone familiar with the subject matter, but those are also the scientists conducting research in the first place. How should we consider the balance between additional time spent on detecting fraud by experts and time spent on traditional research?

Answer: I think that pre- or post-publication review by researchers is a contribution to knowledge-creation, and as such should be recognised and rewarded in ways that are meaningful to researchers, namely in performance and promotion evaluations by their employers, and in research assessment exercises by their funders. That's easier to say than to do for a variety of reasons including, for example, that peer review is usually not a matter of public record. This makes it hard for employers, for example, to validate claims made by researchers about their contributions to peer review. Attempts to address this are not new, and are slowly being trialied by more journals (Nature is trialing transparent peer review — the early results are encouraging, Nature, 2022; Transparent Peer Review at Wiley, Authorea, 2020); other initiatives experiment with new approaches (Review Commons).

It is also worth noting that self-correction is an intrinsic and critical part of the scholarly process which not only identifies fraud and errors, but also identifies clarifications and developments of our understanding. It is fair to note that this process can be slow, perhaps too slow.

I believe that we expose ourselves to integrity risks when we expect peer review alone to do things that are unrealistic, which is discussed in some detail in 'The changing forms and expectations of peer review' (Research Integrity and Peer Review, 2018): "Many have blamed peer review for not properly detecting erroneous research; however, simultaneously, others claim it was never designed to do so." Publishers are creating and implementing enhanced integrity checks to run alongside and augment peer review ('STM Integrity Hub,' STM Association), and to prevent and resolve integrity concerns more effectively. I think there will always be concerns about integrity that readers identify post-publication.

Responses by Dr. Brandon Stell

a. What comments or concerns were raised on PubPeer regarding that paper, and when were those concerns posted? Did PubPeer alert Nature magazine about the issue of the fraudulent images?

The Nature (2006) article by Lesné and coworkers specifically was not discussed on PubPeer before January 2022, by an anonymous user (possibly Schrag):

https://pubpeer.com/publications/8FF7E6996524B73ACB4A9EF5C0AACF

However, similar issues in another of Lesné's papers, in the Journal of Neuroscience, were discussed in 2013:

https://pubpeer.com/publications/841806329CBD4CD2593E83C5E369EA

So a robust investigation of public information could have detected the pattern of issues in Lesné's work almost a decade ago.

PubPeer alerts authors to comments on their articles but it does not alert journals or the authors' institutions unless they have subscribed to our "Dashboard" service (see answers to following questions); to date neither Nature nor the Journal of Neuroscience has subscribed.

As Research Integrity Director at Springer-Nature, witness Chris Graf may be able to describe any private actions taken by Nature after being contacted by Schrag early in 2022. However, it seems to have been the contact from the reporters for *Science* that precipitated the "Editor's Note" in July; Nature readers would have been in the dark about the problems until then.

b. Have other papers had serious concerns raised about the validity of their content on PubPeer? When a similarly significant research misconduct issue is identified on PubPeer, what response do you typically see from the publishing journal, the relevant institution, and the original author?

The PubPeer database contains many thousands of comments highlighting manifest misconduct, typically image manipulations. There are very numerous examples involving the most reputable journals, authors and institutions. Certain authors and laboratories clearly exhibit an historical pattern of such issues.

Authors respond on PubPeer in about 15% of cases, but share convincing original data in only a minority of these. Usually there is no immediate public response from the journal or the institution(s), but that silence does sometimes hide a confidential investigation lasting months or often years. Actions that result when such investigations finally conclude are sometimes considered unsatisfactory - explanations of events can be incomplete, while journals are sometimes fooled into accepting indulgent or manipulated corrections when a retraction may have been warranted. In a recent editorial, the founders of the Retraction Watch project have

argued that the rate of retractions should be at least 20 times higher than it currently is: https://www.nature.com/articles/d41586-022-02071-6

We would add that there is typically no response from funding (e.g. NIH, NSF) or regulatory (e.g. FDA, SEC) agencies.

A typical example of as yet unresolved issues, also drawn from the Alzheimer's field, can be found on this page: https://pubpeer.com/publications/CC6B3B15B1930F4C4E668F112867FC The first post was 10/2020. There are no author responses. The university (not a Dashboard subscriber) became aware at the latest 10/2021 through a user contact. There is a pattern of such issues associated with the senior author, as the approximately 30 results of the following search demonstrate: https://pubpeer.com/search?q=%22Domenico+Pratic%C3%B2%22 To date a single journal has issued a correction, which PubPeer commenters did not judge to be wholly convincing. The university has made no public statement. Thus, for 18 months and counting, only readers consulting PubPeer would be aware of the known problems with this lab and its publications.

c. What processes – if any – are in place to notify institutions or journals when a paper is flagged as potentially fraudulent on PubPeer?

Journals and institutions can be notified in two ways.

Firstly, PubPeer offers a subscription service ("PubPeer Dashboards") that allows journals or institutions to search all commentary relating to them and to receive email alerts when new comments are posted. The revenue from these subscriptions enables the PubPeer Foundation to maintain and develop the PubPeer website. Uptake of this service remains slow, but we have retained all of our subscribers to date. (Annual subscription costs vary depending on comment volume, but are approximately on the order of the processing fee journals charge for publishing a single article.)

Secondly, individual commenters will often contact journals and institutions. Many experienced volunteer analysts such as Elisabeth Bik, Jennifer Byrne and other anonymous PubPeer users frequently do so. However, the consensus among them is that interacting with journals and institutions is particularly time-consuming, frustrating and thankless work. Procedures are still often unclear, opaque, unwelcoming and inefficient, and this strongly discourages volunteers from contacting journals and institutions.

Disclaimer: as mentioned, the PubPeer Foundation raises revenue from subscriptions to its "Dashboard" service.

Appendix II

Additional Material for the Record

Neutron Production and Absorption in Uranium

H. L. ANDERSON, E. FERMI AND LEO SZILARD

Reprinted from THE PHYSICAL REVIEW, Vol. 56, No. 3, August 1, 1939

AUGUST 1, 1939 PHYSICAL REVIEW VOLUME 56
Printed in U. S. A.

Neutron Production and Absorption in Uranium*

H. L. ANDERSON, E. FERMI AND LEO SZILARD
Columbia University, New York, New York

(Received July 3, 1939)

IT has been found[1-3] that there is an abundant emission of neutrons from uranium under the action of slow neutrons, and it is of interest to ascertain whether and to what extent the number of neutrons emitted exceeds the number absorbed.

This question can be investigated by placing a photo-neutron source in the center of a large water tank and comparing, with and without uranium in the water, the number of thermal neutrons present in the water. In the previous experiments of this type[1, 2] it was attempted to have as closely as possible a spherically symmetrical distribution of neutrons. The number of thermal neutrons present in the water was determined by measuring along one radius the neutron density ρ as a function of the distance r from the center, and then calculating $\int r^2 \rho dr$. A difference in favor of uranium of about five percent was reported by von Halban, Joliot and Kovarski.[4]

Since one has to measure a small difference, slight deviations from a spherically symmetrical distribution might give misleading results. The present experiments which are based on the same general principle do not require such symmetry. In order to measure the number of thermal neutrons in the water we filled the tank with a ten-percent solution of $MnSO_4$. The activity induced in manganese is proportional to the number of thermal neutrons present. A physical averaging was performed by stirring the solution before measuring the activity of a sample with an ionization chamber. To obtain an effect of sufficient magnitude, about 200 kg of U_3O_8 was used.

The experimental arrangement is shown in Fig. 1. A photo-neutron source, consisting of about 2 g of radium and 250 g of beryllium was placed in the center of the tank. The geometry was such that practically all neutrons emitted by the source and by the uranium oxide were slowed down and absorbed within the tank. Each irradiation extended over several half-life periods of radiomanganese and the observed activity of the solution was about four times the background of the ionization chamber. Alternating measurements were taken with the cans filled with uranium oxide, and with empty cans of the same dimensions. The activity proved to be about ten percent higher with uranium oxide than without it. This result shows that in our arrangement more neutrons are emitted by uranium than are absorbed by uranium.

In order to find the average number of fast neutrons emitted by uranium for each thermal neutron absorbed by uranium, we have to determine what fraction of the total number of neutrons emitted by the photo-neutron source is, in our experiment, absorbed in the thermal region by uranium. The number of photo-neutrons

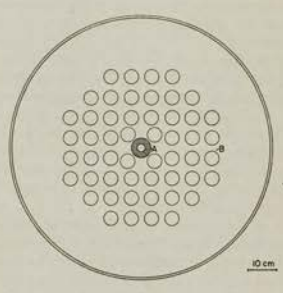

FIG. 1. Horizontal section through center of cylindrical tank which is filled with 540 liters of 10-percent $MnSO_4$ solution. A, Photo-neutron source composed of 2.3 grams of radium and 250 grams of beryllium. B, One of 52 cylindrical cans 5 cm in diameter and 60 cm in height, which are either empty or filled with uranium oxide.

* Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

[1] v. Halban, Joliot and Kovarski, Nature **143**, 470 (1939).
[2] L. Szilard and W. H. Zinn, Phys. Rev. **55**, 799 (1939).
[3] Anderson, Fermi and Hanstein, Phys. Rev. **55**, 797 (1939).
[4] v. Halban, Joliot and Kovarski, Nature **143**, 680 (1939).

emitted by the source is indicated by the activity of the solution in the tank when the irradiation is carried out with empty cans surrounding the source. We obtained a measure of this number by taking into account that in our solution about 20 percent of the neutrons are captured by manganese and the rest by hydrogen. In order to obtain, in the same units, a measure of the number of neutrons absorbed by uranium we proceeded in the following way: A mixture of sand and manganese powder, having the same thermal neutron absorption as uranium oxide replaced the uranium oxide in $\frac{1}{4}$ of the cans which were distributed uniformly among the other uranium oxide-filled cans. After irradiation, all this powder was mixed together, a ten-percent $MnSO_4$ solution was prepared from a sample, and its activity was measured with our ionization chamber.

In this way we found that about 50 percent of the neutrons emitted by the source are absorbed as thermal neutrons by uranium in our arrangement. It follows that, if uranium absorbed only thermal neutrons, the observed ten-percent increase in activity obtained with uranium present would correspond to an average emission of about 1.2 neutrons per thermal neutron absorbed by uranium. This number should be increased, to perhaps 1.5, by taking into account the neutrons which, in our particular arrangement, are absorbed at resonance in the nonthermal region by uranium, without causing neutron emission.

From this result we may conclude that a nuclear chain reaction could be maintained in a system in which neutrons are slowed down without much absorption until they reach thermal energies and are then mostly absorbed by uranium rather than by another element. It remains an open question, however, whether this holds for a system in which hydrogen is used for slowing down the neutrons.

In such a system the absorption of neutrons takes place in three different ways: The neutrons are absorbed at thermal energies, both by hydrogen and uranium, and they are also absorbed by uranium at resonance before they are slowed down to thermal energies. Our result is independent of the ratio of the concentrations of hydrogen and uranium, insofar as it shows that, for thermal neutrons, the ratio of the cross section for neutron production and neutron absorption in uranium is greater than one, and probably about 1.5. What fraction of the neutrons will reach thermal energies without being absorbed will, however, depend on the ratio of the average concentrations of hydrogen and uranium. Since there is an appreciable absorption even far from the center of the resonance band, it follows that the fraction of neutrons absorbed by uranium at resonance will increase with decreasing hydrogen concentration. This has to be taken into account in discussing the possibility of a nuclear chain reaction in a system composed essentially of uranium and hydrogen. A chain reaction would require that more neutrons be produced by uranium than absorbed by uranium and hydrogen together. In our experiment the ratio of the average concentration of hydrogen to uranium atoms was 17 to 1, and in the experiment of von Halban, Joliot and Kovarski this ratio was 70 to 1. At such concentrations the absorption of hydrogen in the thermal region will prevent a chain reaction. By reducing the concentration of hydrogen one would obtain the following effect: On the one hand a larger fraction of those neutrons which reach thermal energies will be absorbed by uranium; on the other hand fewer neutrons reach the thermal region due to an increased absorption by uranium at resonance. Of these two counteracting factors the first is more important for high hydrogen concentrations and the second is more important for low hydrogen concentrations. Starting with high hydrogen concentrations, the ratio of neutron production to total neutron absorption will thus first rise, then pass through a maximum, and, as the hydrogen concentration is decreased, thereafter decrease. We attempted to estimate the quantities involved from the information available about resonance absorption in uranium[5–7] and from the observed net gain of 0.2 in the number of neutrons in our experiment. The effect of the absorption at resonance turns out to be so

[5] Meitner, Hahn and Strassman, Zeits. f. Physik **106**, 249 (1937).
[6] v. Halban, Kovarski and Savitch, Comptes rendus **208**, 1396 (1939).
[7] H. L. Anderson and E. Fermi, Phys. Rev. **55**, 1106 (1939).

large that even at the optimum concentration of hydrogen it is at present quite uncertain whether neutron production will exceed the total neutron absorption. More information concerning the resonance absorption of uranium as well as more accurate measurement of some of the values which enter into our calculation are required before we can conclude whether a chain reaction is possible in mixtures of uranium and water.

We wish to thank Dr. D. W. Stewart, of the Department of Chemistry, and Mr. S. E. Krewer, for advice and assistance in carrying out some of these experiments. We are much indebted to the Eldorado Radium Corporation for enabling us to work with large quantities of uranium oxide in our experiments, and to the Association for Scientific Collaboration for the use of the photo-neutron source and other facilities.

LANCASTER PRESS, INC., LANCASTER, PA.

Uranium: Neutron Production and Absorption

B. FOSTER and E. PERLMUTTER
U.S. House of Representatives, Washington, DC
(Received July 20, 2022)

It has been observed that there is a plentiful emanation of neutrons from uranium (U) under the activity of slow neutrons. It is important to determine whether and how much the quantity of neutrons discharged surpasses the number retained.

Our research team examined this phenomenon by setting a photograph neutron source in the focal point of an enormous tank of water and looking at the quantity of warm neutrons present in the water. The neutrons were examined with uranium added to the tank and without. In the past examinations of this kind, it was endeavored to have as intently as conceivable a roundly even dissemination of neutrons. The quantity of warm neutrons present in the not set in stone by estimating along one span the neutron thickness p as an element of the distance r from the middle, and afterward working out $\int r^2 pdr$. A distinction for uranium of around five percent was accounted for by von Halban, Joliot and Kovarski.

Since one needs to gauge a little distinction, slight deviations from a roundly balanced dissemination could give misdirecting results. The current investigations which depend on a similar general rule don't need such balance. To quantify the quantity of warm neutrons in the water we filled the tank with a 10% arrangement of manganese sulfate (MnSO4). The movement prompted in manganese is corresponding to the quantity of warm neutrons present. We performed a physical averaging by mixing the arrangement prior to estimating the movement of a sample with an ionization chamber. To acquire an impact of adequate extent, around 200 kg of triuranium octoxide (U3O8) was utilized.

A photograph neutron source was put in the focal point of the tank. Around 250 of beryllium (Be) and 2 g of radium (Ra) were added. All neutrons radiated by the source and by the triuranium octoxide were dialed back and retained inside the tank. Every illumination reached out north of a few half-life times of radiomanganese and the noticed movement of the arrangement was multiple times the foundation of the ionization chamber. Rotating estimations were taken with the jars loaded up with triuranium octoxide, and with void jars of similar aspects. When the triuranium octoxide was included, the energy observed in the solution ended up being around 10% higher. This outcome shows that in our course of action a larger number of neutrons are discharged by uranium than are consumed by uranium.

To find the typical number of quick neutrons radiated by uranium for every warm neutron consumed by uranium, we need to figure out what part of the absolute number of neutrons discharged by the photograph neutron source is, in our trial, retained in the warm district by uranium. The quantity of photograph neutrons transmitted by the source is demonstrated by the movement of the arrangement in the tank when the light is done with void jars encompassing the source. We got a proportion of this number by observing that around 20% of the neutrons are caught by manganese (Mn) and the rest by hydrogen (He). To get, in similar units, a proportion of the quantity of neutrons consumed by uranium we continued in the accompanying manner: A combination of sand and manganese powder, having a similar warm neutron retention as triuranium octoxide supplanted the triuranium octoxide in 1/4 of the jars which were disseminated consistently

FAKE

among the other triuranium octoxide-filled jars. After light, everything this powder was combined as one, a 10% MnSO4 arrangement was ready from an example, and its movement was estimated with our ionization chamber.

In this manner we found that around 50% of the neutrons produced by the source are consumed as warm neutrons by uranium in our course of action. Ergo, assuming uranium consumed just warm neutrons, the noticed 10% expansion in movement achieved with uranium present would compare to a typical discharge of around 1.2 neutrons per warm neutron consumed by uranium. This number ought to be expanded, to maybe 1.5, by considering the neutrons which, in our specific course of action, are assimilated at reverberation in the nonthermal district by uranium, without causing neutron emanation.

From this outcome we might reason that an atomic chain response could be kept up with in a framework in which neutrons are dialed back absent a lot of retention until they arrive at warm energies and are then for the most part consumed by uranium as opposed to by another component. It stays an open inquiry, in any case, whether this holds for a framework in which hydrogen is utilized for dialing back the neutrons.

In such a framework the retention of neutrons happens in three unique ways: The neutrons are consumed at warm energies, both by hydrogen and uranium, and they are likewise consumed by uranium at reverberation before they are dialed back to warm energies. Our outcome is autonomous of the proportion of the centralizations of hydrogen and uranium, to the extent that it shows that, for warm neutrons, the proportion of the cross area for neutron creation and neutron assimilation in uranium is more noteworthy than one, and likely around 1.5. What part of the neutrons will arrive at warm energies without being consumed will, be that as it may, rely upon the proportion of the typical centralizations of hydrogen and uranium. Since there is an obvious ingestion even distant from the focal point of the reverberation band, it follows that the negligible part of neutrons consumed by uranium at reverberation will increment with diminishing hydrogen fixation. This must be considered in examining the chance of an atomic chain response in a framework made basically out of uranium and hydrogen. A chain response would expect that a larger number of neutrons be delivered by uranium than consumed by uranium and hydrogen together. In our examination the proportion of the typical convergence of hydrogen to uranium particles was 17 to 1, and in the trial of von Halban, Joliot and Kovarski this proportion was 70 to 1. At such fixations the retention of hydrogen in the warm locale will forestall a chain response. By diminishing the centralization of hydrogen one would get the accompanying impact: On the one hand a bigger part of those neutrons which arrive at warm energies will be consumed by uranium; on the other hand less neutrons arrive at the warm locale because of an expanded retention by uranium at reverberation. Of these two neutralizing factors the first is more significant for high hydrogen fixations and the second is more significant for low hydrogen focuses. Beginning with high hydrogen fixations, the proportion of neutron creation to add up to neutron retention will subsequently first ascent, then go through a greatest, and, as the hydrogen focus is diminished, from that point decline. We endeavored to gauge the amounts required from the data accessible about reverberation retention in uranium and from the noticed net addition of 0.2 in the quantity of neutrons in our trial. The impact of the ingestion at reverberation ends up being enormous to the point that even at the ideal, centralization of hydrogen it is at present very questionable whether neutron creation will surpass the all out neutron retention. More data concerning the reverberation retention of uranium as well as more exact estimation of a portion of the qualities which go into our computation are expected before we can close whether a chain response is conceivable in combinations of uranium and water.

FAKE

We wish to say thanks to Congressman Jay Obernolte, of the 8th District of California, for advice and help with completing a portion of these investigations. We are much obliged to the U.S. Department of Energy for empowering us to work with enormous amounts of triuranium octoxide in our tests, and to the Oak Ridge National Laboratory for the utilization of the photoneutron source and different offices.

FAKE

National Science Foundation • Office of Inspector General
2415 Eisenhower Avenue, Alexandria, VA 22314

July 18, 2022

The Honorable Bill Foster
Chairman, Subcommittee on Investigations & Oversight
House Committee on Science, Space & Technology
United States House of Representatives

The Honorable Jay Obernolte
Ranking Member, Subcommittee on Investigations & Oversight
House Committee on Science, Space & Technology
United States House of Representatives

Dear Chairman Foster and Ranking Member Obernolte:

I write regarding the hearing planned for July 20, 2022, titled "Paper Mills and Research Misconduct: Facing the Challenges of Scientific Publishing," which will discuss challenges in securing scientific literature from fraudulent academic papers.

Subcommittee staff provided us with a fake physics paper created by using a paper generator web site. The fake paper was based on a famous physics paper from 1939. At the request of Subcommittee staff, our investigators ran the fake paper through the commercially available plagiarism detection software we use in the course of our work. The software did not find any plagiarism. We confirmed the original paper is available on the internet by running it through the same software, which identified it.

Next, we used a document-to-document comparison function in our software to compare the fake paper with the original. It found a 35% overlap. We also used freely available software to directly compare the two documents, and it found 32% overlap. This high overlap would raise flags for investigators and reviewers.

In summary:

- If our office received the fake paper as part of an allegation, but did not know about the original 1939 paper, we would not have identified a problem by using plagiarism detection software alone. However, it would be obvious to a human reviewer that the paper was fake based on the non-sensical replacements made by the paper generator site. For example, the site replaced:
    - "photon" with "photograph" multiple times
    - "spherically symmetrical distribution" with "roundly balanced dissemination"

- "slowed down and absorbed" with "dialed back and retained"
- "absorbed at resonance" with "assimilated at reverberation"
- "Our result is independent of the ratio of the concentrations of hydrogen and uranium" with "Our outcome is autonomous of the proportion of the centralizations of hydrogen and uranium"

- If our office received the fake paper as part of an allegation and the reviewer identified the potential source, both our commercially available software and the freely available software would find enough overlap to warrant further investigation. In this case, we would open an inquiry, conduct a more thorough review of the paper, contact the author to ask about the overlap, and look at other documents written by the author. If the author's response to our questions did not dispel the allegation, we would refer an investigation to the grantee and provide the evidence we gathered.

Please feel free to contact me at 703-292-7100 or my chief of staff, Lisa Vonder Haar, at lvonderh@nsf.gov if you have any questions.

Sincerely,

Allison C. Lerner

Allison C. Lerner
Inspector General