



HAL
open science

Fraud, exaggerations, and ghost authorship: questioning the reliability of the scholarly record

Frédérique Bordignon, Guillaume Cabanac, Cyril Labbé, Maud Bernisson

► To cite this version:

Frédérique Bordignon, Guillaume Cabanac, Cyril Labbé, Maud Bernisson. Fraud, exaggerations, and ghost authorship: questioning the reliability of the scholarly record. *Ingenius*, 2023. hal-04094363

HAL Id: hal-04094363

<https://enpc.hal.science/hal-04094363>

Submitted on 11 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fraud, exaggerations, and ghost authorship: questioning the reliability of the scholarly record

The credibility of science is regularly undermined by highly prominent cases of fraud, but it also suffers from more discreet examples of bad practices and misconduct, which can be just as harmful. Concerned scientists focus on studying academic production, with projects designed to identify anomalies and errors in the literature, in order to gain insights into the causes and study possible correction mechanisms. The authors of this article are contributors to one such project, called [NanoBubbles](#).

By **Frédérique Bordignon**^{1,2}, **Guillaume Cabanac**^{3,4}, **Cyril Labbé**⁵ and **Maud Bernisson**⁶

¹ Ecole des Ponts, 77455 Marne-la-Vallée, France

² LISIS, INRAE, Univ Gustave Eiffel, CNRS, 77455 Marne-la-Vallée, France

³ IRIT UMR 5505 CNRS, Université Toulouse 3 – Paul Sabatier, France

⁴ Institut universitaire de France (IUF), Paris, France

⁵ Université Grenoble-Alpes, CNRS, Grenoble INP, Laboratoire d'Informatique de Grenoble, Grenoble, France

⁶ Radboud University, Nijmegen, The Netherlands

The most high-profile cases are those that go beyond the confines of the scientific community, since they are reported or revealed by the media, and sometimes even end up in court. Interested readers may visit the [RetractionWatch](#) website, which lists and explains such cases. Here, we will cite a few cases to demonstrate the diversity of problems that specialists describe in terms of falsification and fabrication of data, plagiarism, which are all practices that violate scientific integrity. In 2021, a neuroscientist revealed that images used fourteen years earlier as scientific evidence in an article reporting the validation of a hypothesis to explain Alzheimer's disease had been falsified. This post-publication report undermined a research area that had been considered promising until that time, and had received millions of dollars in investment. Another case concerns a prominent Dutch psychologist who confessed in 2011 that he had forged experiments and therefore completely fabricated data, producing findings that were spectacular – but falsified. And yet, these unreliable conclusions had been widely reported in the media, thereby spreading false ideas such as, among others, the fact that employers consider job applicants to be more competent when they have a masculine voice.

Experts and forgers [\[1\]](#)

When it comes to plagiarism, it is now commonly said to have little effect on the validity of the science. However, it strongly impacts the plagiarized authors, who must often fight to have the damage recognized. Plagiarism refers to both stealing ideas and stealing research and has always existed. It has been reported that in the second century AD, Ptolemy took credit for astronomical measurements made 300 years earlier by Hipparchus in Rhodes, whereas he claimed to have made them on the Egyptian coast. The consequences of scientific fraud vary in significance: it may unfortunately go unnoticed, but can also lead to the stigmatization of the forgers or to the mutilation or [death of patients](#). It can also influence public acceptance of health policies (as with the promotion of hydroxychloroquine which was never proved to be effective against Covid-19, or the false causal

relationship between the measles vaccine and autism in 1998, which led to a sharp drop in vaccination in the United Kingdom)[2].

While such fraud is sometimes revealed by whistleblowers who have witnessed this cheating, other cases are brought to light by sleuths who scrutinize scientific publications. Some even specialize in this area and have turned it into a professional activity, like microbiologist [Elisabeth Bik](#) who detects manipulated images, such as histological sections.

Fraud detectors

Given the magnitude of the task of tracking fraud and misconduct, Guillaume Cabanac and Cyril Labbé are developing IT tools to scan the 6 million scientific papers published per year. The nine detectors of the Problematic Paper Screener (PPS) have flagged 12,000 problematic papers to date. Some of the "papers" are entirely meaningless (automatically generated with SCIfgen or Mathgen). Biology articles mention erroneous DNA sequences (which *Seek&Blastn* detects) or use contaminated cell lines. "Tortured phrases" [3] result from copying–paraphrasing–pasting (producing "nucleic corrosive" instead of "nucleic acid", and "counterfeit consciousness" for "artificial intelligence", for example). Tortured phrases were found in two thirds of the problematic papers listed by the PPS. Forgers use paraphrasing software to mask their plagiarism, even if it means destroying the meaning of the original text. The PPS also detects articles "with feet of clay," drawing on unreliable references which are retracted or criticized elsewhere.

Scientists concerned about such problematic publications reassess the articles flagged by the PPS and post their reports on [PubPeer](#). Discussions between authors and publishers then take place, sometimes leading to corrections or retractions. By automating the search for flaws in the 120 million papers published to date, the PPS facilitates the reassessment of the most questionable results – an essential activity in order to accumulate reliable knowledge.

Exaggerations and false promises

Far from the extreme cases mentioned above, scientific papers may suffer from exaggerations that mislead readers. Researchers tend to unconsciously oversell their findings and exaggerate their potential benefits, by using overly positive expressions, as shown by Frédérique Bordignon and her colleagues at the very beginning of the Covid crisis in a study [4] on the sharp increase in the use of the adjectives *effective*, *promising*, and *novel* (which the journal [ACS Catalysis had already banned](#) from its titles and abstracts in 2017) to describe obtained results. These rhetorical manipulations are common, especially in abstracts, which seek to catch readers' attention to entice them to read the entire study. A more critical distortion of reality, referred to as spinning, occurs when authors go beyond linguistic exaggeration and present only their best results: those that are the most appealing, and will create hype. Unfortunately, these promises then spill over into press releases without nuance and, in turn, into mainstream media articles, spreading false hopes and misinformation.

Ghost authors

Maud Bernisson has also analyzed how the pharmaceutical industry uses scientific communications to disseminate marketing messages about the benefits of a newly–patented drug, or by introducing a new concept for the benefit of a drug to increase its sales (for example, the diagnosis of "pseudoaddiction" was invented to show that patients who are suffering should have their dose of opioid medication increased (study sample: one patient [5])). Medical education and communication agencies (MECCs) play the role of ghostwriters, writing papers whose conclusions strongly support the company's interests, then ask honorary authors to sign the papers in their name [6].

Lastly, among the extreme cases of ghostwriting, we can cite [paper mills](#), which offer ready-to-publish articles for hundreds or even thousands of euros. Unfortunately, the papers are of poor

quality, often based on fabricated data, faked or plagiarized images, and may lack meaning, which are all clues into their dubious origin (which the PPS tracks).

Virtual authors

The capabilities of artificial intelligence (AI) tools now make it possible to generate images and texts that seem plausible, but have a questionable scientific basis. While they do not outright reject such progress, some publishers are already taking steps to update their editorial policies to prohibit or [acknowledge](#) the use of AI tools for writing scientific papers. Most of the time, these editorial policies do not authorize the inclusion of ChatGPT as an author, since the author of a publication must be able to answer for their work.

The question of who can, must, and may not be considered an author is continuously raised in new ways, whether it's a matter of [adding one's cat or hamster as a co-author](#), or using the pseudonym [Camille Nous](#) for protest purposes – yet solutions must be found in order to improve and maintain the integrity of the scientific literature. This is precisely the aim of the researchers working on the [NanoBubbles](#) ERC synergy project funded by the EU and carried out in France and the Netherlands by an interdisciplinary team of physicists, historians, sociologists, philosophers, linguists, and computer scientists.

[1] Bessy, C., & Chateauraynaud, F. (2014). *Experts et faussaires: Pour une sociologie de la perception* (2e éd., augmentée d'une postface). Éditions Pétra.

[2] See the documentary “[La fabrique de l'ignorance](#)” on Arte TV.

[3] Cabanac, G., Labbé, C., & Magazinov, A. (2021). *Tortured phrases: A dubious writing style emerging in science. Evidence of critical issues affecting established journals* (arXiv:2107.06751). arXiv. <https://doi.org/10.48550/arXiv.2107.06751>

[4] Bordignon, F., Ermakova, L., & Noel, M. (2021). Over-promotion and caution in abstracts of preprints during the COVID-19 crisis. *Learned Publishing*, leap.1411. <https://doi.org/10.1002/leap.1411>

[5] Michaels, D., & Michaels, D. (2020). *The Triumph of Doubt: Dark Money and the Science of Deception*. Oxford University Press.

[6] Sismondo, S., & Doucet, M. (2010). Publication Ethics and the Ghost Management of Medical Publication. *Bioethics*, 24(6), 273–283. <https://doi.org/10.1111/j.1467-8519.2008.01702.x>