Peer Reviewed | In&Vertebrates

The file drawer effect – a long-lasting issue in the sciences

Bastien S Lemaire[1,2], Raphaël Lenoble[1], Mirko Zanon[1], Thibaud Jacquel[2], Giorgio Vallortigara[1]

[1] Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy
[2] inandsight.science, Bordeaux, France

https://doi.org/10.52732/KUWG1044

QUESTION

What are the causes and consequences of the file drawer effect?

ABSTRACT

Most of the scientific outputs produced by researchers are inaccessible since they are not published in scientific journals: they remain in the researchers' drawers, forming what we call the Dark Science. This is a long-standing issue in research, creating a misleading view of the scientific facts. Contrary to the current literature overfed with positive findings, the Dark Science is nurtured with null findings, replications, flawed experimental designs and other research outputs. Publishers, researchers, institutions and funders all play an important role in the accumulation of those unpublished works, but it is only once we understand the reasons and the benefits of publishing all the scientific findings that we can collectively act to solve the Dark Science problem. In this article, we discuss the causes and consequences of the Dark Science expansion, arguing that science and scientists would benefit from getting all their findings to the light of publication.

Keywords: file drawer effect, dark science, publication bias, null findings, replications, flawed designs

For correspondence:
bastien.lemaire@unitn.it

Peer Reviews: The peer review reports are published alongside the article.

Received: 30 Nov. 2021
Revised: 20 Jan. 2022
Published: 27 Jan. 2022

THE DARKSCIENCE, A.K.A THE FILE DRAWER EFFECT

Science and scientists work best when they have full access to literature. Unfortunately, most of the scientific outputs produced are inaccessible. Half of the published experiments are still blocked behind paywalls [1] and an even bigger number is tucked away in file drawers, composing what we call the '*Dark Science*'.

While the growth of preprint platforms and open access scientific journals – where the published content is freely accessible – has begun to solve the former issue (+25 % of open access articles since 2000 [1]), the Dark Science problem remains relatively unhandled.

The Dark Science problem, commonly called the file drawer problem, refers to a culture of seeking and publishing only certain kinds of results, mostly those that reach the magic p=0.05. In 1979 [2], Rosenthal described that scientists' file drawers were filled with 95% of non-significant results while the published literature was mainly composed of positive findings – with a probabilistic 5% of false positives (Type I error). In a recent report, Allen and Melher [3] confirmed Rosenthal's claim. The current literature contains from 5 to 20% of null findings, while the literature obtained from pre-registered reports – when the methods and the hypotheses of an experiment are published before acquiring the data – has 4 to 10 times more. This clearly suggests that the traditional literature is a misleading representation of the scientific facts.

While it is still unclear how many experiments fill the Dark Science every year, we know that it is widespread in the sciences [4][5] and that it has most probably increased with time since the proportion of positive findings grew by over 15% from 1990 to 2007 [6].

In the following sections of this article, we identify (1) the scientific outputs that are more likely to be stored in file drawers rather than being published in top scientific journals, (2) the causes and consequences of the Dark Science expansion and (3) the benefits of bringing the Dark Science back to the sight of science and scientists.

DARK SCIENCE, CONTENT, CAUSES AND CONSEQUENCES

Nowadays, articles stuck in the Dark Science are mostly null (statistically non-significant) findings, replications and flawed designed experiments.

Null findings are not all similar [7]. In fact, we can categorize them into two main types: inconclusive and conclusive findings [8]. Inconclusive findings often emanate from studies with small sample sizes (such as sometimes in field studies) that lack sufficient statistical power and are hardly interpretable [9][10]. Conclusive findings, instead, can be interpreted with more certainty.

Replications are also less likely to be published in top scientific journals. This is particularly true for replications yielding no effect or confirming established results. However, replications demonstrating the opposite effect of established results seem to escape more easily from the Dark Science [11].

Flawed designs are poorly documented too [11]. Often, scientists fail at designing proper experimental designs to answer their hypothesis, but also fail at disseminating failures.

Null findings, replications and flawed designs are common in research. In fact, they probably constitute the majority of the research results and are often discussed informally in laboratories during lab meetings given their importance into the making of scientific discoveries. However, the rest of the scientific community does not benefit from them as they remain unpublish.

In 2014, null findings were 40% less likely to see the light of publication and in 60% of the case, they were not even written up [4]. It appears difficult to come up with one main reason explaining this approach, still, we can point out how publishers, researchers and funders (either passively or actively), all participate in this phenomenon.

Publishers

Most of the time, scientific journals seek breakthroughs. This leads to publication biases, which occur, according to Dickersin [12], when "publication of study results is based on the direction or significance of the findings". Publication biases are globally observed in all the scientific disciplines and countries [6]. As stated in the previous section [3], the scientific literature is only composed of 5 to 20% of null findings, while it should approximate 60%. One reason is that null findings (as well as replications and flawed designs) do not fill the knowledge gap. At least, not at first. Therefore, their impact is difficult to evaluate at the moment of publication. According to Johnson and Dickersin [7], a less common but more worrying explanation to publication bias is that editors direct their article selection towards eye-catching content for readers; in the worst case, intentionally for financial purposes (in 2017 scientific publishing annual market represented $10 billion [13]).

Besides, the publishing costs indirectly affect publication bias. Researchers with limited funds prioritize their work that will be most likely to be published: novel, positive and therefore more attractive to most journals.

However, more and more journals and publishing platforms are seeking reliability rather than novelty. They willingly encourage scientists to submit their null findings, replications and 'failures' to disseminate them to the broadest audience (see In&Sight, PLOS one, PeerJ, JOTE, etc.). Publishers and journals appear to not be the only ones at fault in the Dark Science expansion, researchers play a major role too.

Researchers

For Johnson and Dickersin [7], the main reason why null findings are not published is that researchers fail to submit them. This is somehow confirmed by Franco and collaborators [4] who demonstrated that 60% of null findings are not even written up although most scientists agree that they are worth being published [11][14]. There are multiple reasons that lead researchers to keep part of their research unpublished. As described above, the cost to publish is one of them.

Researchers have limited funds, and it costs on average $1500 to publish an article in an open access journal. Therefore, they prioritize some work at the expense of others. Time is limited as well, and it appears to be the main reason why researchers do not publish all of their findings [14]. Considering that publishing is a long process – ranging from a couple of months to one year to publish an article [15] – and that publishers are mostly interested in novel findings, it is not surprising that much (if not most) research ends up in the Dark Science. For replications and flawed designs experiments, the same reasons probably apply [16]. Specifically related to flawed designs, or what we could call failures, researchers could feel embarrassed with their work and, therefore, more likely to put it aside. In a recent article, Peterson and Panofsky [17] interviewed 60 members of the Board of Reviewing Editors for the journal Science and showed that the motives for replication differ between fields and produces different cultures of replication. In some cases, replications are impossible as the associated costs are too high. In other cases, replications are performed to extend knowledge on top of the replicated study and are not intended to be published. Going back to null findings, another reason that pop-up from Echevarria's study [14] is that null findings are less cited than positive results (although this is only true in specific fields [18] and therefore, are not submitted for publication to journals. This is a very interesting point as it brings the '*incentives to publish*' component on the table: how researchers are evaluated by their institutions and funding agencies is of fundamental relevance; these last are therefore responsible too in the Dark Science expansion.

Institutions/Funders

The growing competition for research funding and academic positions, combined with the use of specific bibliometrics (h-index, impact factor) led researchers into a 'publish or perish' culture: publish as many articles as possible in high-impact factor journals [19][20]. In Life Sciences, it is still widely believed that you need to author a paper in high-impact factor journals such as Cell, Nature or Science (CNS) to secure a faculty position and receive job offers [21][22]. As those journals are highly competitive and seeking breakthroughs, it strongly influences what researchers focus on and, very likely, prevent them from spending time on publishing other findings, which instead end up in the Dark Science [23]. Until institutions and funders use new/different metrics to select the next scientists' generation, researchers will continue to look for high-impact factor journals, with publishers promoting theirs to attract researchers – a never-ending loop.

The pressures to publish are indeed intense and inevitably influence researchers but it is also good to remember that publication metrics do not necessarily correlate with the percentage of job offers received [22] and that next-generation metrics in line with open science practices are arising [24][25]. By changing the way institutions and funders evaluate researchers, the incentives to publish can change toward better practices; practices that could be beneficial to science and the community.

As stated at the beginning of this section, it is difficult to come up with one reason to explain the Dark Science expansion which appears inevitable in today's research culture. Publishers, researchers and funders/institutions all play a role and the entire ecosystem requires changes.

PUBLISHING THE DARK SCIENCE, WHAT ARE THE BENEFITS?

We argue that many findings of the Dark Science such as null findings, replications and flawed designs, would be beneficial to science if published and addressed adequately.

Null but conclusive findings have the power to dissuade researchers from unfruitful avenues. Contrarily, null but inconclusive findings are hardly interpretable. However, they are useful in laying down scientific questions that remain to be tackled by the scientific community. Surprisingly or not, once published and grouped in systematic reviews, those underpowered studies can be lifesaving. This is well illustrated by the Stroke Unit Trialists' collaborative systematic review who gathered a series of null and underpowered studies [26] which once together demonstrated that having a stroke unit care reduces the odds of death by 17% (95% confidence interval 4-29%). Similar examples have been described in cancer, cardiovascular and thyroid-related research, demonstrating the importance of publishing such studies [27][28][29]. Replications validate or invalidate published findings and make science reliable. They are therefore crucial to science. However, at the moment less than 25% of the replicated studies confirm the original results [30][31] and many papers fail to disseminate enough methodological details to allow replications [32]. However, true (or diagnostic) replications are not always necessary to validate or invalidate previously described results. Indeed, the convergence of evidence through triangulations of experiments can act as indirect replications and, at the same time, push the boundaries of knowledge. Although they may appear embarrassing, flawed designs are important too. Those studies should be used to learn from each other's mistakes. In today's publishing process, flawed designs are mostly identified by peer reviewers prior to publication. Unfortunately, in most cases, it leads to the rejection of the article and not to the discussion of the detected problem. Yes, science is often flawed, but it is time we embrace it since this is certainly the only activity which allow us to be actually wrong (at least 5% of the times). Some scientific investigations find their way out of the Dark Science and the traditional publishers' hands; they compose the so-called grey literature. They are various in form (government reports, theses and dissertations, conference papers and even non-written resources such as posters and infographics) and are essential resources in systematic reviews [33]. However, the grey literature is hard to find and not peer reviewed, making its use even more challenging.

Overall, publishing the Dark Science can only lead to a more reliable, collaborative and complete science. Beyond doubts, publishing the Dark Science would save time, effort and funds. The amount of resources saved is difficult to evaluate, but Glazsiou and Chalmers [34] did get an approximation of the funds lost every year in medical research only. They estimated that 85 % of the research was wasted, which corresponds to $170 billion per year; a number that should send a chill up your spine. As for saving funds, the amount of time we could save is incalculable. However, we can count the number of hours scientists spend reviewing articles that are then rejected. It represents 15 million hours every year [35]. Necessary, but wasted as the articles are buried in the Dark Science, potentially with precious information to convey to the scientific community.

To make reliable science and discoveries, scientists need access to all the information. However, this brings up an important limitation that should be tackled. As the number of scientific articles increases every year, it is more and more difficult for scientists to stay up to date and find relevant literature for their research. Publishing the Dark Science would most likely double the number of articles published every year, making the search for relevant content even harder. Chu and Evans [36] pointed out that excessive publication can lead to scientific stagnation rather than progress. In a new light of publishing all findings, we would like to stress our opposite position. Considering the current publication biases, scientific progress is limited: many findings are not replicable, and the statistical information to properly evaluate a correct power analysis is often missing. It follows that conclusive and strong scientific claims are almost impossible to make, leading to the well-known replicability crisis. This is a strong sign of stagnation. It is evident that publishing all the scientific findings will increase researchers' time to find and evaluate the literature. Still, this process will result in more substantial and unbiased materials to construct research on, which in return will undoubtedly lead to scientific progress. Publishing the Dark Science would make the search for relevant content even harder, as it would double the number of articles published every year. However, the internet has offered tools previously inconceivable. Those tools, and the use we make of them, have the possibility to change the way we select, verify, publish and search science. After decades of keeping our research in the drawers, it is time to unleash our Dark Science with open eyes.

References

[1] Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, Farley A, West J, Haustein S (2018) **The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles** *PeerJ,* 6:e4375
https://doi.org/10.7717/peerj.4375

[2] Rosenthal R (1979) **The file drawer problem and tolerance for null results.** *Psychological Bulletin,* 86:638-641
https://doi.org/10.1037/0033-2909.86.3.638

[3] Allen C, Mehler DMA (2019) **Correction: Open science challenges, benefits and tips in early career and beyond** *PLOS Biology,* 17:e3000587

https://doi.org/10.1371/journal.pbio.3000587
[4] Franco A, Malhotra N, Simonovits G (2014) **Publication bias in the social sciences: Unlocking the file drawer** *Science*, 345:1502-1505
https://doi.org/10.1126/science.1255484
[5] Brown AW, Mehta TS, Allison DB (2017) **Publication Bias in Science** *Oxford Handbooks Online*,
https://doi.org/10.1093/oxfordhb/9780190497620.013.10
[6] Fanelli D (2011) **Negative results are disappearing from most disciplines and countries** *Scientometrics*, 90:891-904
https://doi.org/10.1007/s11192-011-0494-7
[7] Johnson RT, Dickersin K (2007) **Publication bias against negative results from clinical trials: three of the seven deadly sins** *Nature Clinical Practice Neurology*, 3:590-591
https://doi.org/10.1038/ncpneuro0618
[8] Sandercock P (2011) **Negative Results: Why Do they Need to be Published?** *International Journal of Stroke*, 7:32-33
https://doi.org/10.1111/j.1747-4949.2011.00723.x
[9] Mlinarić A, Horvat M, Šupak Smolčić V (2017) **Dealing with the positive publication bias: Why you should really publish your negative results** *Biochemia Medica*, 27
https://doi.org/10.11613/bm.2017.030201
[1 0] Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR (2013) **Power failure: why small sample size undermines the reliability of neuroscience** *Nature Reviews Neuroscience*, 14:365-376
https://doi.org/10.1038/nrn3475
[1 1] Tsou A, Schickore J, Sugimoto CR (2014) **Unpublishable research: examining and organizing the 'file drawer'** *Learned Publishing*, 27:253-8
https://doi.org/10.1087/20140404
[1 2] Dickersin K (1990) **The existence of publication bias and risk factors for its occurrence.** *JAMA*, 263:1385-9
http://www.ncbi.nlm.nih.gov/pubmed/2406472
[1 3] Johnson R, Watkinson A, Mabe M (2018) **The STM Report: an overview of scientific and scholarly publishing.**
https://www.stm-assoc.org/2018_10_04_STM_Report_2018.pdf
[1 4] Echevarría L, Malerba A, Arechavala-Gomeza V (2021) **Researcher's Perceptions on Publishing "Negative" Results and Open Access** *Nucleic Acid Therapeutics*, 31:185-189
https://doi.org/10.1089/nat.2020.0865
[1 5] Runde BJ (2021) **Time to publish? Turnaround times, acceptance rates, and impact factors of journals in fisheries science** *PLOS ONE*, 16:e0257841
https://doi.org/10.1371/journal.pone.0257841
[1 6] Kulkarni S **Why are replication studies so rarely published?** *Editage Insights*,
https://doi.org/10.34193/ei-a-6230

[17] Peterson D, Panofsky A (2021) **Self-correction in science: The diagnostic and integrative motives for replication** *Social Studies of Science*, 51:583-605
https://doi.org/10.1177/03063127211005551
[18] Fanelli D (2012) **Positive results receive more citations, but only in some disciplines** *Scientometrics*, 94:701-709
https://doi.org/10.1007/s11192-012-0757-y
[19] De Rond M, Miller AN (2005) **Publish or Perish** *Journal of Management Inquiry*, 14:321-329
https://doi.org/10.1177/1056492605276850
[20] McGrail MR, Rickard CM, Jones R (2006) **Publish or perish: a systematic review of interventions to increase academic publication rates** *Higher Education Research & Development*, 25:19-35
https://doi.org/10.1080/07294360500453053
[21] McKiernan EC, Schimanski LA, Muñoz Nieves C, Matthias L, Niles MT, Alperin JP (2019) **Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations** *eLife*, 8
https://doi.org/10.7554/elife.47338
[22] Fernandes JD, Sarabipour S, Smith CT, Niemi NM, Jadavji NM, Kozik AJ, Holehouse AS, Pejaver V, Symmons O, Bisson Filho AW, Haage A (2020) **A survey-based analysis of the academic job market** *eLife*, 9
https://doi.org/10.7554/elife.54097
[23] Fanelli D (2010) **Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data** *PLoS ONE*, 5:e10271
https://doi.org/10.1371/journal.pone.0010271
[24] Wilsdon J, Bar-Ilan J, Frodeman R, Lex E, Peters I, Wouters P (2017) **Next-generation metrics: responsible metrics and evaluation for open science**
https://doi.org/10.2777/337729
[25] DORA **San Francisco Declaration on Research Assessment** *DORA*,
https://sfdora.org/
[26] Collaboration SUT **Organised inpatient (stroke unit) care for stroke: network meta-analysis**
https://doi.org/10.1002/14651858.cd000197.pub4
[27] Simes RJ (1986) **Publication bias: the case for an international registry of clinical trials.** *Journal of Clinical Oncology*, 4:1529-1541
https://doi.org/10.1200/jco.1986.4.10.1529
[28] Alderson P (2004) **Absence of evidence is not evidence of absence** *BMJ*, 328:476-477
https://doi.org/10.1136/bmj.328.7438.476
[29] Rennie D (1997) **When authorship fails. A proposal to make contributors accountable** *JAMA: The Journal of the American Medical Association*, 278:579-585
https://doi.org/10.1001/jama.278.7.579

[30] Prinz F, Schlange T, Asadullah K (2011) **Believe it or not: how much can we rely on published data on potential drug targets?** *Nature Reviews Drug Discovery*, 10:712-712
https://doi.org/10.1038/nrd3439-c1
[31] Begley CG, Ellis LM (2012) **Raise standards for preclinical cancer research** *Nature*, 483:531-533
https://doi.org/10.1038/483531a
[32] Carp J (2012) **The secret lives of experiments: Methods reporting in the fMRI literature** *NeuroImage*, 63:289-300
https://doi.org/10.1016/j.neuroimage.2012.07.004
[33] Paez A (2017) **Gray literature: An important resource in systematic reviews** *Journal of Evidence-Based Medicine*, 10:233-240
https://doi.org/10.1111/jebm.12266
[34] Chalmers I, Glasziou P (2009) **Avoidable waste in the production and reporting of research evidence** *The Lancet*, 374:86-89
https://doi.org/10.1016/s0140-6736(09)60329-9
[35] The AJE Team **Peer Review: How We Found 15 Million Hours of Lost Time**
https://www.aje.com/arc/peer-review-process-15-million-hours-lost-time/
[36] Chu JSG, Evans JA (2021) **Slowed canonical progress in large fields of science** *Proceedings of the National Academy of Sciences*, 118:e2021636118
https://doi.org/10.1073/pnas.2021636118