'It could be effective…': Uncertainty and over-promotion in the abstracts of COVID-19 preprints

A defining feature of the impact of the COVID-19 pandemic on scholarly communication has been the brief and intense surge in the production of preprints. This has had significant impacts on the ways in which new research findings have been reported and communicated more broadly and the role played by abstracts in highlighting the meaning and value of new research. Based on a study of the language deployed in the abstracts of recently published COVID-19 preprints, **Frédérique Bordignon, Liana Ermakova** *and* **Marianne Noël***, argue two defining features of these abstracts are over-promotion and hedging, a deliberate ambiguity that suggests authors should pay greater attention to what they seek to communicate in their abstracts.*

The coronavirus outbreak has '[broken the mold](#)' of traditional publishing and has changed how scientists communicate. In particular, since the beginning of the COVID-19 pandemic, the number of deposits of preprints surged, thus challenging the small teams in charge of screening them. [We have been studying abstracts](#) for several years and [we decided to investigate](#) whether this crisis and the ensuing scientific and economic competition have changed how academics write and present their research in abstracts.

We know that, even under normal circumstances, researchers tend to exaggerate the benefit of their research and use promotional discourse in order to publish, advance their career, or obtain funding. The abstract is known to be a promotional genre with the overuse of positive terms and optimistic phrases intended to convince the reader (e.g., 'it could be effective…', 'our findings suggest promising effects on…', 'this present study could provide a novel insight into…'). Over-promotion in abstracts aims at catching the reader's attention and inviting them to read further.

[In a recent study](#), we found that with the need to establish priority and stand out from the unprecedented volume of preprints, authors made greater use of positive adjectives. But, being aware of the uncertainty surrounding their results, they also toned down their discourse, inviting readers to be cautious.


We reached these seemingly contradictory conclusions *via* a lexical analysis of preprint abstracts urgently produced in response to the COVID-19 crisis (thanks to the [corpus](#) built by Nicholas Fraser in 2020 and still regularly updated by himself and Bianca Kramer) compared to a control corpus of abstracts produced before the pandemic.

Our study was initially inspired by that of Vinkers et al. on the use of positive and negative words in PubMed abstracts published between 1974 and 2014. They showed the tendency to over-promote results was increasing, leading them to conclude that researchers 'assume that results and their implications have to be exaggerated and overstated in order to get published'. Building on this study, we in turn analysed a corpus of 23,957 preprints abstracts deposited on 7 different servers (SSRN, arXiv, medRxiv, bioRxiv, Research Square, Preprints.org, and ChemRxiv) and found that, since the beginning of the pandemic, authors have more frequently used words indicating the validity of their approach (*effective, efficient, robust*) and their potential impact in the fight against the virus (*promising*) and they have also attempted to show the innovative nature of their work by highlighting novelty (*novel, unique*) in this competitive context.

> scientists themselves must be aware that the excessive use of positive words which they try to balance with many hedge words is just a band-aid solution and does not work in favour of an informative abstract

As far back as the early 1990s, Goodman had already suggested that *novel* was in danger of becoming no more than the synonym of *new*. Our study of COVID-19 preprints abstracts might throw light on the weak signal that *effective* will soon supplant *novel* to qualify the latest medical advances. The use of this word reflects a stronger focus on proof of concept (whose purpose is to verify that some concept or theory has the potential of being used), rather than on the priority of the announced results. Fortunately, if authors sought to be more impactful than during the pre-pandemic period, they did not massively resort to using the words *groundbreaking*, *miraculous* or *stunning*, which more clearly imply exaggeration and overstatement.

In contrast to over-promotion, hedging is a procedure of argumentation allowing writers to tone down their statements, and thereby gain community acceptance and keep critics away. We compiled a list of hedge words (e.g., *possible, quite, might, perhaps, will, may, suggest*) and counted them in the corpus. We showed that there is a strong increase in hedge words (the most frequent of which are the modal verbs *can* and *may*) in the COVID-19 preprints abstracts. They counterbalance the excessive use of positive words and thus invite the readers, who go probably beyond the 'usual' audience, to be cautious with the obtained results, illustrating the balance that authors have to achieve between promoting their results and appealing for caution. Our corpus of 'crisis literature' is teeming with such examples: '*This imaging modality <u>could</u> be <u>effective</u> for the diagnosis of enterocolitis associated with COVID-19*.'
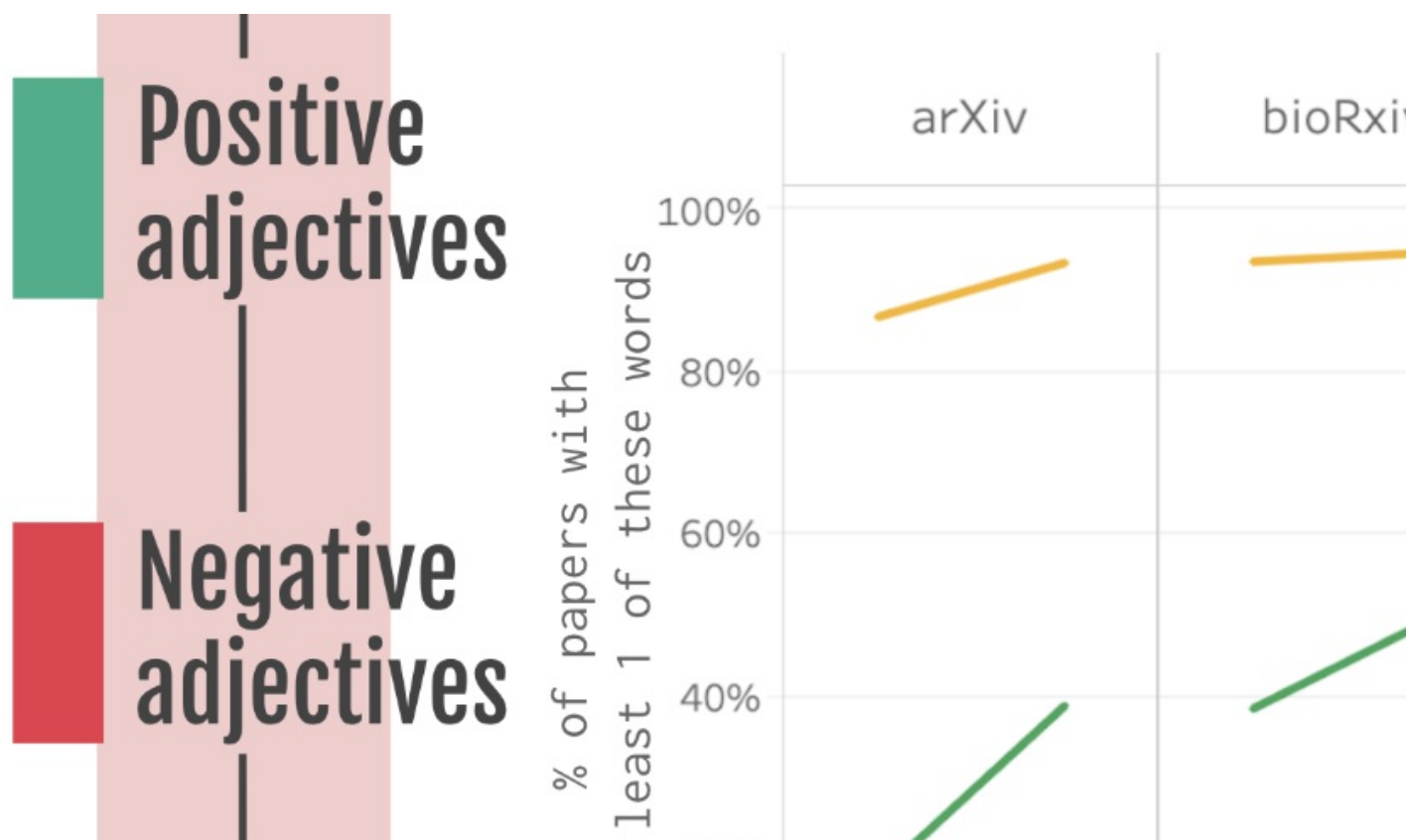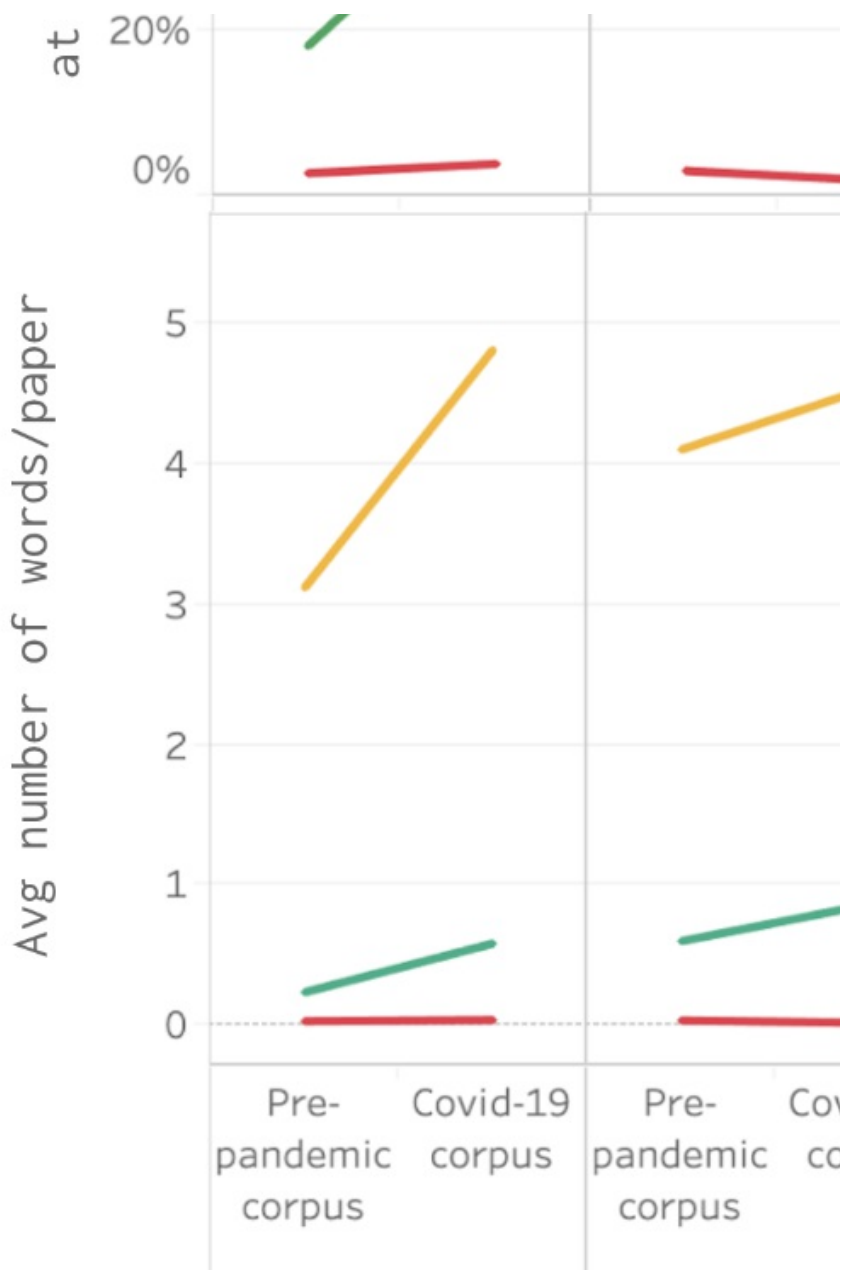
Results synthesis per preprint server: percentage of paper abstracts with positive/negative adjectives and hedge words, average number of positive/negative adjectives and hedge words per abstract. Comparison before and during the COVID-19 crisis

Publishers have tried to curb exaggeration and spinning, considered as a form of misconduct, in that they are ways of distorting science reporting without actually lying. Publishers attempt to establish guidelines for formatting the presentation of results (e.g., the CONSORT statement for Randomized Controlled Trials, the PRISMA statement for abstracts of systematic reviews) or even to prohibit the inappropriate use of exaggerating words (e.g., *ACS Catalysis* journal). But, scientists themselves must be aware that the excessive use of positive words which they try to balance with many hedge words is just a band-aid solution and does not work in favour of an informative abstract. Science hype in preprint abstracts can have detrimental consequences, especially in times of crisis. Abstracts are often the only part of the work that will be read and therefore the over-promotion will raise false expectations that will not be counterbalanced by the reading of the full-text. The combination of promise and uncertainty conveyed by a large number of preprints made available to all may have contributed to the 'infodemic' denounced by the WHO in June 2020. Our recommendation is that authors (including ourselves) should take the time to carefully write their abstracts and ponder on the use of words to avoid being inaccurate. Abstracts are a window into the long form of the scientific article and inaccuracies in abstract risk discrediting the scientific article itself.

This post draws on the authors' article, Over-promotion and caution in abstracts of preprints during the COVID-19 crisis, published in Learned Publishing.

Note: This article gives the views of the author, and not the position of the LSE Impact Blog, nor of the London School of Economics. Please review our comments policy if you have any concerns on posting a comment below

Image Credit: Adapted from J J Ying via Unsplash.