

**Main Manuscript for**

Big team science as a response to urgent societal developments.

Nicholas A. Coles, Joao Francisco Goes Braga Takayanagi

1Department of Psychology, University of Florida; Gainesville, FL.

\* Nicholas A. Coles

**Email:**  [ncoles@ufl.edu](mailto:ncoles@ufl.edu)

**Author Contributions:** Conceptualization: NAC. Methodology: NAC. Investigation: NAC, JFGBT. Visualization: NAC. Funding acquisition: NAC. Project administration: NAC, JFGBT. Supervision: NAC. Writing – original draft: NAC. Writing – review & editing: NAC, JFGBT.

**Competing Interest Statement:** Authors declare that they have no competing interests.

**Classification:** Social Sciences | Psychological and Cognitive Sciences

**Keywords:** big team science, collaboration, urgent developments, impact, speed

**This file includes:**

Main Text

Figures 1

**Abstract**

When responding to urgent societal developments – e.g., the rise of terrorism, pandemics, and disruptive technologies – researchers often collaborate en masse. The present work evaluates the impact of these big team science initiatives via an examination of a near-population of 2,150,491 scholarly articles published on terrorism, COVID-19, and ChatGPT. Results indicate that big team science is (a) an increasingly common response to urgent societal developments, that (b) is associated with increased impact on science, policy, and news, with (c) currently little-to-no concomitant delay in published insights. As researchers prepare responses to both existing and future urgent challenges, results suggest they can increase their societal impact via big team science.

**Main Text**

**Introduction**

The first quarter of the 21st century has presented society with numerous developments that have demanded quick and urgent action, including some of the deadliest terrorist attacks in modern history, a global pandemic, and the ongoing rise of AI technologies. The present work focuses on researchers’ responses to these urgent developments, evaluating the role of a recent trend: big team science (1).

The present work focuses on research published by relatively small (1 – 4 co-authors) vs. big research teams (up to >100+ co-authors) during three significant societal developments: one historical (the September 11, 2001 terrorism attacks), one recent (the COVID-19 pandemic), and one ongoing (the public release of the AI tool, ChatGPT). A near-population of research on these topics is compiled via OpenAlex: an openly available catalogue of scientific articles (2). To evaluate impact, we leverage OpenAlex and Altmetrics (3) to examine how often articles are discussed in other scholarly publications, news outlets, and policy documents. Last, we evaluate the speed of researchers’ responses by examining article publication dates.

**Results**

Results are organized around three main insights.

***Researchers are increasingly responding to urgent societal developments with big team science***

Consistent with broader trends in science (4), big team science is an increasingly common response to urgent societal developments. As an example, consider researchers’ response to the COVID-19 pandemic to an event that occurred nearly two decades prior: the September 11 terrorism attacks.

Researchers have used the term ‘terrorism’ in 322,878 articles – 89.81% of which were published after the September 11 attacks. This research has been overwhelmingly published by relatively small teams. Approximately 93.66% of articles have fewer than 5 co-authors, and less than 1% of articles have 10 or more co-authors. Figure 1 focuses on a range of team sizes that consistently had more than 30 papers to evaluate: collaboration with up to 40 collaborators. Intriguingly, such efforts often yielded more impact. However, there were also preliminary signs of diminishing returns among the largest collaborations, as well as reductions in speed (Figure 1).

Nearly two decades later, collaborative infrastructure had evolved substantially when researchers began responding to another urgent development: the COVID-19 pandemic. Compared to the September 11 attacks, researchers’ responses were substantially more collaborative. The proportion of published articles with fewer than 5 co-authors decreased 27.28 percentage points (now 66.38%), the proportion of articles with 5 to 9 co-authors increased four-fold (now 23.96%), and the proportion of articles with 10 or more co-authors increased eight-fold (now 8.71%), all *χ2*(1, *N* = 2114551) > 24821, *p* < .001.

***When responding to the COVID-19 pandemic, bigger teams had more impact on science, the news, and policy – with little-to-no speed tradeoff***

In research on COVID-19, articles published by 1 - 4 co-authors (*N* = 1,208,365) were mentioned, on average, in 5.33 scholarly works (*SD* = 34.90), 1.52 news outlets (*SD* = 15.33), and 0.10 policy documents (*SD* = 0.76; Figure 1). Comparatively, articles with 10 – 19 co-authors (*N* = 141,051) were mentioned 3.21 times more in scholarly works (*M* = 17.12, *SD* = 90.79), 2.68 times more in news outlets (*M* = 4.07, *SD* = 36.95), and 1.80 times more in policy documents (*M* = 0.18, *SD* = 1.29), all *W* > 14730910329, *p* < .001. As an even more extreme comparison, consider COVID-19 articles with 100 or more co-authors (*N* = 1,631). Compared to small team efforts, these ultra-large collaborations were mentioned 17.68 times more in scholarly works (*M* = 94.20, *SD* = 381.81), 22.22 times more in news outlets (*M* = 33.71, *SD* = 139.38), and 9.97 times more in policy documents (*M* = 0.99, *SD* = 4.97), all *W* > 120102939, *p* < .001.

Of note, big teams did not usually encounter substantial time delays (Figure 1). In fact, they often published scientific insights quicker than their small team counterparts. On average, small teams of 1 - 4 co-authors published their insights 802.48 days after the World Health Organization declared a pandemic (*SD* = 408.51). Larger teams – such as those containing 10 – 19 (*M* = 798.59 days, SD = 427.24) or 40 – 49 authors (*M* = 733.63 days, *SD* = 413.54) – typically published at a slightly quicker pace, all *W* > 1849189519, *p* < .001. However, some exceptions exist. For example, compared to small teams of 1 – 4 co-authors, the largest teams of 100+ co-authors tended to publish their insights 34 days later (*M* = 836.97 days, *SD* = 408, *W* = 938646379, *p* < .001).

***Looking forward: How are scientists collaboratively tackling the emergence of transformative AI technologies?***

Almost two years after responding to COVID-19, scientists encountered yet another urgent global development: the public release of the AI chatbot, ChatGPT. Although certainly not the only technology on the market, ChatGPT is often considered one of the most influential – receiving credit for accelerating an AI boom whose economic, social, and environmental impacts are rapidly evolving but still unknown (5). Furthermore, the phrase “ChatGPT” is virtually nonexistent in the scientific literature before the product’s release (November, 2022). Like research on terrorism and COVID-19, this provides a unique methodological opportunity to evaluate scientists’ response time.

Similar to the COVID-19 pandemic, research on ChatGPT (*N* = 35,940) signals a shift towards big teams – with 4.45% of published articles having 10 or more co-authors. There are too few papers to reliably evaluate the impact of their largest collaborative responses; articles with 40 - 49 authors (*N* = 19), 60 - 69 authors (*N* = 6), and 80 - 89 authors (*N* = 1) are still rare. Nonetheless, the pattern observed in collaborations up to those sizes suggest that history may very well repeat itself: big teams will have an outsized impact on the development and understanding of emerging AI technology.

Once again, we first consider the norm: papers published by smaller teams of 1 – 4 co-authors (*N* = 26,324; 73.24%). In research on ChatGPT, these relatively small teams have received an average 4.66 mentions in other scholarly work (*SD* = 27.05) 1.48 mentions in the news (*SD* = 17.96) and 0.03 mentions in policy documents (*SD* = 0.43). Comparatively, the 1,392 articles with 10 - 19 co-authors have received 2.70 times more mentions in scholarly works (*M* = 12.58, *SD* = 69.79, *W* = 14151318, *p* < .001), 2.21 times more mentions in news (*M* = 3.27, *SD* = 28.53, *W* = 2290250, *p* = .006), and a similar number of mentions in policy documents (*M* = 0.03, *SD* = 0.30, *W* = 2394508, *p* = .93). The 106 articles with 20 – 29 co-authors have fared even better, receiving 7.99 times more mentions in scholarly works (*M* = 37.27, *SD* = 169.64, *W* = 1063268, *p* < .001), 1.32 times more mentions in news (*M* = 1.96, *SD* = 5.51, *W* = 162342, *p* = .01), and 3.99 times more mentions in policy documents (*M* = 0.13, *SD* = 0.60, *W* = 179715, *p* = .04). Once again, these larger collaborations are achieving increased impact with minimal speed tradeoffs. On average, small teams of 1 – 4 co-authors published their insights 5.72 days quicker than teams of 5 – 9 (*W* = 102811076, *p* < .001), 19.20 days quicker than teams of 10 – 19 (*W* = 17140707, *p* < .001), and 19.13 days *slower* than teams of 30 – 39 co-authors (although this last difference is not statistically significant, *W* = 582757, *p* = .38).

**Discussion**

In summary, the present work highlights the growing role of big team science as a response to urgent societal developments. The analysis of research following the September 11 attacks, COVID-19 pandemic, and release of transformative AI technologies like ChatGPT reveals a trend: during urgent societal developments, researchers are increasingly collaborating en masse to produce timely and high-impact insights. Indeed, an influential paper describing GPT-4 was co-authored by a team of over 200 authors (6) – as was a paper describing an influential competitor, DeepSeek (7). Despite the impact of ultra-large collaborations, however, they remain relatively uncommon due to barriers related to training (8), evaluation (4), and publication norms (9, 10). As we prepare for future urgent challenges, we must adapt our existing frameworks to not just accommodate but also encourage scientists to leverage the power of mass collaboration to maximize societal impact.

**Materials and Methods**

Scholarly articles were compiled between September and October 2024 via OpenAlex: an open-source bibliographic catalogue of scientific papers. Records were removed if they (a) had duplicative DOI’s, (b) did not contain authorship meta-data, or (c) were published before an event signaled the societal development. For the latter, terrorism papers had to be published after the date of the attacks (September 11, 2001); COVID-19 records had to be published after the first cases emerged in November 2019; and ChatGPT records had to published after the public release of the tool (November 30, 2022). This left *N* = 289,986 terrorism, *N* = 1,824,565 COVID-19, and *N* = 35,940 ChatGPT papers.

Team size was extracted using OpenAlex authorship meta-data. Multiple operationalizations of “big team science” were examined: publications with (a) 5 – 9, (b) 10 – 19 authors, (c) 20 – 29 authors, (d) 30 – 39 authors, (e) 40 – 49 authors, (f) 50 – 59 authors, (g) 60 – 69 authors, (h) 70 – 79 authors, (i) 80 – 89 authors, (j) 90 – 99 authors, and (k) 100+ authors. These big team efforts are compared to publications with relatively smaller teams of 1 – 4 co-authors.

To index impact, we evaluate how often these articles are discussed in other scholarly publications, news outlets, and policy documents. For the first outcome, we examined citation meta-data from OpenAlex. For the latter two outcomes, we use data from Altmetric: a proprietary service that tracks online engagement with scholarly works.

For publication speed, OpenAlex publication dates were used to determine how many days had passed since an event signaled an urgent societal development. For terrorism papers, we used September 11, 2001. For COVID-19 papers, we used the date the World Health Organization declared a pandemic (March 11, 2020). For ChatGPT data, we used the date the tool was made publicly available (November 30, 2022).

**Acknowledgments**

This work was supported by funding from the John Templeton Foundation (#62295) and National Science Foundation (#2235070).

**References**

Sample references:

1. B. Nogrady, Hyperauthorship: the publishing challenges for ‘big team' science. *Nature* *615*, 175-177 (2023).
2. J. Priem, H. Piwowar, R. Orr, OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv*: 2205.01833 [cs.DL] (2022).
3. J. L. Ortega, Reliability and accuracy of altmetric providers: a comparison among Altmetric. com, PlumX and Crossref Event Data. *Scientometrics*, *116*, 2123-2138 (2018).
4. N. A. Coles, J. K. Hamlin, L. L. Sullivan, T. H. Parker, D. Altschul, Build up big-team science. *Nature*, *601*, 505-507 (2022).
5. K. Weise, C. Metz, N. Grant, M. Isaac. Inside the A.I. arms race that changed Silicon Valley forever. *New York Times* (2023). Available at https://www.nytimes.com/2023/12/05/technology/ai-chatgpt-google-meta.html.
6. OpenAI, J. Achiam, S. Adler, et al., GPT-4 Technical Report. *arXiv*: 2303.08774 [cs.CL] (2024).
7. DeepSeek-AI, D. Guo, D. Yang, et al., DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning. *arXiv*: 2501.12948 [cs.CL] (2025).
8. S. M. Fiore, C. Gabelica, T. J. Wiltshire, D. Stokols, “Training to be a (team) scientist. Strategies for team science success” in Strategies for Team Science Success (Springer, 2019), pp. 421-444.
9. N. A. Coles, L. M. DeBruine, F. Azevedo, H. A. Baumgartner, M. C. Frank, ‘Big team’ science challenges us to reconsider authorship. *Nature Human Behaviour*, *7*, 665-667 (2023).
10. N. A. Coles, E. R. Tenney, J. M. Chin, J. C. Friedrich, R. E. O’Dea, A. O Holcombe, Team scientists should normalize disagreement. *Science*, *384*, 1076-1077 (2024).

**Figures**

A graph of different colored lines

AI-generated content may be incorrect.

**Figure 1. Impact and response speed of research conducted on terrorism, COVID-19, and ChatGPT.** Relationship between team-size (x-axis), response speed (green line), and mentions in scholarly articles (pink line), news outputs (blue line), policy documents (orange line). Bigger teams have also exhibited little-to-no difference in response speed (green), showing small improvements in response time in some cases and small decreases in response time in other cases. In the graph below, dots indicate average percentiles. Solid dark lines represent relationships estimated via generalized additive models, and colored bands represent 95% confidence intervals.