

The replication crisis and the need to change the policy of scientific publication¹

Hadi Samadi (Islamic Azad University of Tehran, hadi.samadi@yahoo.com)

ARTICLE INFO

Article History

Received: 2021/1/25

Accepted: 2021/2/20

Key Words:

Reproducibility Crisis,

Statistics,

P Value,

Cognitive bias,

Open Science

ABSTRACT

Over the past decade, we have increasingly realized that when we try to replicate the data which are published in influential scientific journals, the initial results are not obtained. In other words, the stated findings of the articles which are published in the experimental sciences, including medicine and psychology, are generally invalid. This phenomenon has led to a crisis in the experimental sciences, known in 2012 as the "replication crisis." In the present article, first, the crisis is introduced and a brief account of the opinions of the scientists about the severity of the crisis is produced. Then, five possible causes for its occurrence are introduced. The claim of the article is that the most important cause for the crisis of reproducibility of scientific results of academic articles should be sought in the wrong policies of the academic institute or funding bodies of putting pressure on scientists to publish articles (publish or perish). In the final section, some solutions, corresponding to the five causes which are identified, to alleviate the crisis are presented. Two of these solutions, namely, expanding the domain of the "open science" and secondly, reducing the role of articles in assessing the competence of researchers and academics are introduced as the main strategies to alleviate the crisis.

1. The first draft of this article was presented as a lecture at the Iranian Academy of Medical Sciences. I took advantage of the criticisms of the participants. The second draft of the article was read and annotated by Dr. Ali Paya. I also thank him for his really useful comments.

بحران تکرارپذیری و ضرورت تغییر در سیاست‌گذاری چاپ مقالات علمی^۱

هادی صمدی (دانشگاه آزاد اسلامی تهران؛ hadi.samadi@yahoo.com)

اطلاعات مقاله

تاریخ دریافت: ۱۳۹۹/۱۱/۶

تاریخ پذیرش: ۱۳۹۹/۱۲/۲

واژگان کلیدی:

بحران تکرارپذیری،

آمار،

ارزش پی،

سوداری شناختی،

علم باز

چکیده

طی دهه گذشته به نحوی روزافزون دریافته‌ایم که وقتی درصد تکرار داده‌های نشر یافته در مقالات معتبر علمی هستیم، نتایج اولیه به دست نمی‌آیند. به عبارتی بسیاری از مقالات نشر یافته در علوم آزمایشی، از جمله پزشکی و روان‌شناسی نامعتبر هستند. این پدیده به بحرانی در علوم تجربی منجر شده است که در ۲۰۱۲ به «بحران تکرارپذیری» معروف شد. در مقاله حاضر ابتدا بحران معرفی، و نظر دانشمندان در مورد شدت آن بیان شده است. سپس پنج علت احتمالی برای بروز آن معرفی شده‌اند. دعوی مقاله آن است که مهم‌ترین علت را باید در سیاست‌گذاری‌های ناکارآمد فشار بر چاپ مقالات جستجو کرد. در بخش بعد نیز متناسب با پنج علت راهکارهایی برای تخفیف بحران معرفی شده است. در بخش پایانی گام برداشتن به سوی «علم باز» و کاستن از نقش مقالات در ارزیابی استادان و پژوهشگران به‌عنوان دو راهکار اصلی معرفی شده است.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

۱. متن اولیه این مقاله در یک سخنرانی در فرهنگستان علوم پزشکی عرضه شده است. از نقدهای شرکت‌کنندگان در جلسه در اصلاح متن بهره بردم و ممنونم. نسخه اولیه مقاله را جناب دکتر علی پایا مطالعه فرمودند و نکات ارزشمندی را یادآوری کردند. از زحمات ایشان نیز سپاسگزارم.

مقدمه: بحران تکرارپذیری چیست؟

از زمان شیوع بیماری کرونا به دفعات شاهد آن بودیم که شواهدی عرضه شده است مبنی بر اینکه داروی «الف» (یا شیوه درمانی «الف») در کاهش علائم بیماری یا درمان آن مؤثر است؛ اما با گذشت تنها چند روز شواهد دیگری عرضه شده‌اند مبنی بر نادرستی شواهد نخست. اگر بپذیریم که پارادایم رایج در پزشکی امروز، پارادایمی است معروف به «پزشکی مبتنی بر شواهد»^۱ پرسش این خواهد بود که چه هنگامی موجه خواهد بود که شواهد له یا علیه یک نظریه (مثلاً اینکه داروی «الف» علائم کرونا را تخفیف می‌دهد) را در پزشکی مبنای اقدامات درمانی قرار دهیم. ساده‌ترین پاسخ به این پرسش این است: «هرگاه شواهد محصول روش‌های پژوهش موثق و مبتنی بر آمار باشند».

پیش فرض این پاسخ آن است که اگر شواهد محصول روش‌های پژوهش موثق مبتنی بر آمار باشند، در تکرارهای بعدی همان پژوهش، کم‌وبیش همان نتایج را خواهیم داشت و بنابراین به نحوی مشابه، بیماران نیز، همانند گروه مورد آزمایش، درمان خواهند شد. اگر این شواهد، فرضیه مورد آزمون را در گام نخست قبول یا رد کرده باشند، انتظار بر آن است که در گام بعدی نیز چنین کنند. در گفتار حاضر ابتدا نشان خواهیم داد پیش فرض ظاهراً مقبول ذکر شده با چالش‌های زیادی مواجه شده است؛ چالشی که در سال ۲۰۱۲ «بحران تکرارپذیری» نام گرفت. در گام دوم به علل این بحران پرداخته و در پایان به راه‌های برون‌رفت از آن اشاره خواهد شد.

به‌طور سنتی فیلسوفان زیادی، به‌ویژه در سنت قاره‌ای، بوده‌اند که شأن ممتازی را که علم در جهان امروز دارد به چالش کشیده‌اند (به‌عنوان نمونه هیدگر را ببینید که با فهم پوزیتیویستی رایج در زمان خود از علم مخالف بود).^۲ فیلسوفانی نیز در سنت تحلیلی بوده‌اند که خنثی بودن مشاهدات را به پرسش گرفتند و اعلام کردند که همه مشاهدات وابسته‌اند به نظریه‌ها و پیش فرض‌های نظری مشاهده‌گر (به‌عنوان نمونه پوپر، هسنس،^۳ فایرماند،^۴ و کوون^۵ را ببینید). فیلسوفان دیگری از همین سنت نیز

عنیت آزمایش را به پرسش گرفته‌اند (به‌عنوان نمونه کالینز^۶ و پیکرینک^۷ را ببینید). اما این بار تفاوت در آنجا بود که چالش از سوی خود دانشمندان پیش رو نهاده شده بود و در واقع برخی یافته‌های تجربی گواهی بودند بر اینکه عموم داده‌هایی که گمان داشتیم موثق‌اند چندان هم موثق نبوده‌اند.

مطابق فهم عرفی دانشمندان وقتی یک فرضیه آماری در سطح آلفای پنج‌صدم به چاپ می‌رسد به معنای آن است که اگر ۱۰۰ بار آزمایش یادشده تکرار شود، انتظار آن است که در حدود ۹۵ بار همان نتایج قبلی به دست آید (در ادامه خواهیم دید که یک فیلسوف علم، الکساندر برد، این فهم عرفی را به چالش کشیده است). اما تکرار آزمایش‌های یادشده در این مقالات نشریافته نتایج کاملاً متفاوتی را نشان می‌داد. در یکی از پژوهش‌ها که در ادامه معرفی می‌شود نشان داده شد که ۶۵ درصد مقالات نتیجه‌ای متفاوت از پژوهش چاپ‌شده داشتند. ۶۵ درصد به میزان قابل توجهی از عدد مورد انتظار، یعنی پنج درصد بیشتر است. پرسش این بود که به چه میزان می‌توان به یافته‌های علمی موجود در نشریات معتبر اعتماد کرد؟ پرسش دیگر اینکه علل به‌وجود آمدن این بحران کدامند؟ در ادامه تلاش می‌شود دلایل اصلی بحران معرفی شوند اما قبل از آن اشاره مختصر به نحوه آشکار شدن بحران خواهیم کرد.

آگاهی از وجود بحران

تجربه مشترکی میان عموم دانشجویان رشته‌هایی که دروس آزمایشگاهی داشته‌اند وجود دارد: اینکه بارها تمامی دستورالعمل‌های توصیه‌شده برای به ثمر رسیدن یک آزمایش را رعایت کرده‌اند اما به نتیجه مطلوب نرسیده‌اند. تا بیست سال پیش کمتر کسی گمان آن داشت که این تجربه به بحرانی در علم بینجامد. سال ۲۰۰۵ مقاله‌ای توسط یوانیدس، پژوهشگر علوم پزشکی، به چاپ رسید با این عنوان که «چرا عمده یافته‌های پژوهشی چاپ‌شده نادرست‌اند».^۸ این بار نه دانشجویان کم‌تجربه، بلکه پژوهشگری صاحب‌نام مدعی شده بود که اگر

5. Kuhn, T.S., The Structure of Scientific Revolutions, 1962.

6. Collins, H., 1985.

7. Pickering, A., 1987.

8. Ioannidis, John P. A. (August 1, 2005).

1. Evidence-Based Medicine Working Group (1992).

2. Rouse, Joseph. Heideggerian philosophy of science, 1998.

3. Feyerabend, P.K., 1959.

4. Feyerabend, P.K., 1959.

دیگر رشته‌ها نشود؛ درحالی‌که بحث یوانیدس عمدتاً متوجه پزشکی بود.

وقتی سال ۲۰۱۲ نام «بحران تکرارپذیری» وضع شد، وجود مشکلی اساسی به رسمیت شناخته شد. حال پرسش این بود که اولاً چقدر مشکل جدی است و ثانیاً علل آن کدامند؟ اگر می‌توانستیم به این دو پرسش پاسخی درخور دهیم آنگاه راه برون‌رفت نیز می‌توانست پیدا شود.

پرسش نخست این بود که «مشکل چقدر جدی است». می‌توان به نحوی پیشینی درباره‌ی شدت یک بحران و عوارض آن سخن گفت؛ اما دانشمندان ترجیح می‌دهند پاسخ چنین پرسش‌هایی را به نحوی تجربی به‌دست آورند. بنابراین، همه‌پرسی بزرگی از خود دانشمندان صورت گرفت و نتیجه در ویدئویی که در نیچر ۲۰۱۶ منتشر شد^۲ که همین پرسش را از ۱۵۷۶ دانشمند برتر جهان می‌پرسید. تنها سه درصد گفتند بحرانی وجود ندارد. هفت درصد ابراز بی‌اطلاعی کردند و نود درصد وجود مشکل را پذیرفتند.

این نود درصد به دو زیرگروه تقسیم می‌شدند: ۳۸ درصد مشکل را خفیف دانستند و ۵۲ درصد مشکل را حاد ارزیابی کردند. بنابراین از منظر اکثریت دانشمندان مشکل جدی ارزیابی شد. در همان نظرخواهی از دانشمندان درباره‌ی علل بحران نیز نظرسنجی شد. دو علت بیش از دیگران مورد توجه بودند: مشکلات پیرامون آمار و روش تحقیق، و فشارهای اجتماعی که بر دانشمندان برای چاپ سریع مقالات وارد می‌شود. اما در ادامه تلاش می‌شود افزون بر شرح این دو، به دیگر عوامل دخیل نیز اشاره شود.

علل بحران تکرارپذیری

علت بحران چیست؟ در پاسخ به این پرسش اختلاف نظرها اساسی‌تر بود. در اینجا تلاش می‌شود انواع علل را در پنج دسته قرار دهیم:

۱. علت را باید در تفاوت در شرایط آزمایش و به‌ویژه در تفاوت در مدل‌های آزمایش جست.

درصد آزمون مجدد یافته‌های نشریافته در مجلات معتبر برآیند درخواست یافت که آن یافته‌ها کاذب‌اند. بنابراین کوتاهی از جانب دانشجویان نبوده است، تقصیری اگر هست متوجه خود مقالات، و به‌ویژه در روش‌شناسی و فرایند چاپ مقالات است.

همان‌طورکه اشاره شد دهه‌ها بود که فیلسوفانی در نقد عینی انگاشتن یافته‌های پژوهش علمی سخن گفته بودند؛ اما این‌بار دانشمندی نامدار، نقدی بسیار تند بر اعتبار علم وارد می‌کرد و نه فیلسوفی که بتوان با زدن انگ ناآشنایی با علم نقدهای او را نادیده گرفت. بنابراین بحث وارد مرحله‌ای جدید شد. با بالا گرفتن بحث، دانشگاه ویرجینیا در آمریکا بودجه‌ای اختصاص داد تا دعای‌ای از این سنخ بررسی شوند و بیش از این اعتبار علم خدشه بر ندارد.

به‌عنوان نمونه، سه مجله معتبر روان‌شناسی انتخاب، و ۱۰۰ مقاله که در سال ۲۰۰۸ در این مجله‌ها چاپ شده بودند به صورت نمونه‌گیری تصادفی از میان آنها گزینش شدند و تلاش شد آزمایش‌ها به نحوی حتی الامکان مشابه مقاله اصلی تکرار شوند. یافته‌های جدید در سال ۲۰۱۱ چاپ شد^۱ و نتیجه بسیار مایوس‌کننده بود. تنها یافته‌های ۳۵ مقاله تکرار شدند؛ آن هم با آلفایی بسیار بیش از آنچه در مقالات اصلی ادعا شده بود و بنابراین نشان داده شد احتمال خطای نوع اول، یعنی چاپ شدن مقاله‌ای که مدعی وجود رابطه‌ای بین پدیده‌هاست درحالی‌که در واقع چنین رابطه‌ای وجود ندارد، بسیار بالاتر از دعوی مقالات است (آلفا احتمال خطای نوع اول یا مثبت کاذب است).

اگر ارزش پی را به‌طور متعارف در حد پنج‌صدم در نظر بگیریم و بپذیریم مقالات چاپ‌شده نیز همگی آلفایی کمتر از پنج‌صدم داشته‌اند، مطابق فهم عرفی دانشمندان، بدان معناست که انتظار داشتیم حداقل ۹۵ یافته تکرار شوند. اما تنها ۳۵ مقاله تکرار شده بود؛ یعنی حق با یوانیدس بود (در ادامه خواهیم دید که برد نشان می‌دهد عدد مورد انتظار باید ۶۷ باشد و نه ۹۵، اما حتی در آن صورت نیز ۳۵ بسیار کمتر از حد انتظار است). سال متعاقب نشر این پژوهش آماج حملات متوجه روان‌شناسی بود؛ زیرا عموماً ترجیح داده می‌شد که کسی متوجه این بحران در

2. Nature Video (28 May 2016).

1. Weir, Kristen. 2011.

راهکارهای بدیلی برای پژوهش‌های علمی یافت.^۲ اما در مقابل، حتی در حیطه هوشناسی نیز نظرات لورنتس به چالش کشیده شده است و عده‌ای معتقدند حتی سیستم‌های آب و هوایی نیز به حساسیتی که لورنتس در نظر گرفته نیستند و با در نظر گرفتن چند عامل پیش‌بینی ممکن است^۳ و اگر خطایی در پیش‌بینی می‌بینید به سراغ خطاهای اندازه‌گیری بروید.^۴

فرض کنید پدیده مورد مطالعه ابتدا به کروناس و فرض پژوهش این است که دو عامل «الف» و «ب» در ابتدا به کرونا دخالت دارند. این اعتراض که عوامل متعددی بر ابتدا به کرونا مؤثرند سخنی درست، اما بی‌فایده است؛ زیرا راهکاری عرضه نمی‌کند، مگر آنکه از عوامل دیگری نام‌برده شود که در این صورت با فرضیه‌ای کاملاً جدید روبه‌رو هستیم و تلاش برای آزمودن آن دیگر ربطی به آزمایش قبلی محسوب نمی‌شود و نباید آن را تلاشی در تکرار آزمایش قبلی در نظر گرفت که به شکست انجامیده است. نکته دیگر این است که افزودن هر عامل جدیدی به مدل آزمایش قبلی خطای اندازه‌گیری تازه‌ای را نیز وارد مدل می‌کند که ممکن است خود به تکرارناپذیری‌های بعدی در آزمون فرضیه جدید منجر شود.

ب. علت دوم بحران تکرارپذیری را می‌توان به تفاوت‌های میان آزمودنی‌ها نسبت داد. در فحوی این سخن نهفته است که به‌عنوان نمونه اگر داروی «الف» در درمان بیماران کرونایی چینی مؤثر بوده اما در درمان بیماران آلمانی مؤثر نبوده است اختلاف دو حالت را باید در تفاوت چینی‌ها و آلمانی‌ها جست. این سخن حقیقی از حقیقت را داراست: اینکه تفاوت‌هایی در ساختمان بدنی و جسمی افراد موجود در نژادهای مختلف، بین زن و مرد، بین پیر و جوان و غیره وجود دارد و این تفاوت می‌تواند به اختلاف در نتایج دو آزمایش بینجامد.

در واقع توجه به همین اختلاف‌ها بوده است که اساس پزشکی شخصی شده را شکل می‌دهد.^۵ اما به دو دلیل انداختن تمامی بار بحران تکرارپذیری بر شانه‌های تفاوت فردی خطاست. نخست

۲. علت را باید در اختلاف آزمودنی‌های موجود در گروه‌های آزمایشی پی گرفت.

۳. مشکل در تفاوت میان آزمایشگران است.

۴. مشکل در روش‌های آماری مورد استفاده است.

۵. مشکل در زمینه اجتماعی تولید مقالات علمی است.

به نظر می‌رسد مجموعه‌ای از همه این موارد باعث انحراف‌های بزرگ از یافته‌های اولیه می‌شود و باید متواضعانه پذیرفت در همه ابعاد مشکلاتی وجود دارد؛ اما خواهیم گفت که چرا پنجمی را باید جدی‌تر گرفت.

بررسی یک به یک

الف. وقتی می‌گوییم مشکل در شرایط آزمایش است؛ بدین معناست که شرایط آزمایش به درستی تکرار نشده‌اند. این خود به دو معنا می‌تواند باشد. اینکه یا با همان عواملی به‌عنوان متغیر مستقل وارد مدل آزمایش شده‌اند نتایج تکرار نشده است (که خود دو حالت دارد: به دلیل عدم کنترل متغیرهای مداخله‌گر یا با وجود کنترل متغیرهای مداخله‌گر) یا با افزودن یا کم کردن تعداد متغیرهای مستقل.

بی‌شمار از عوامل می‌توانند بر وقوع هر پدیده مورد بررسی اثر گذارند. هرچه پدیده مورد مطالعه آشوبناک‌تر باشد تعداد این عوامل بیشتر می‌شود تا جایی که در شکل افراطی آن لورنتس اثرات پر زدن پروانه را در مکانی بسیار دور بر وقوع طوفانی در مکانی نزدیک منتفی نمی‌داند.^۱ حتی اگر به لحاظ هستی‌شناسی سخن لورنتس درست باشد، به لحاظ معرفت‌شناسی پذیرش آن مانعی بر نظریه‌پردازی و آزمون‌های علمی در رشته‌هایی مانند پزشکی و روان‌شناسی می‌شود. بنابراین، به‌طور متعارف با وجود آنکه هر پژوهشگری آگاه است که افزون بر متغیرهای مستقلی که او در نظر گرفته است عوامل دیگری نیز می‌توانند بر پدیده مورد بررسی اثر گذارند؛ اما به لحاظ عملی چاره‌ای جز این نمی‌بیند که تنها عوامل محدودی را در مدل آزمایشی وارد کند تا امکان بررسی فراهم باشد.

برخی معتقدند همین انحراف از واقعیت علت اصلی بحران تکرارپذیری است و باید به سراغ نظریه پیچیدگی رفت و

1. Lorenz, Edward N. (March 1963).

2. Wallot, Sebastian; Kelty-Stephen, Damian G. (2018).

3. Orrell, David; Smith, Leonard; Barkmeijer, Jan; Palmer, Tim (2001).

4. Orrell, David (2012).

5. Geoffrey S. Ginsburg, Huntington Ph.D Willard 2009.

پس از چاپ مقاله‌ای در یک مجله معتبر پژوهش‌های بعدی در آن زمینه با موانعی برای گرفتن بودجه، انجام و سپس انتشار مواجه می‌شوند. اگر پژوهشی نشان داده است که داروی «الف» در درمان بیماری «ب» مؤثر است و یافته‌های پژوهش در مجله‌ای معتبر مانند لنتس به چاپ رسیده باشد، عموم دانشمندان علوم پزشکی با آن یافته به‌عنوان یک فکت برخورد می‌کنند و کسی درصدد آزمون مجدد آن برنمی‌آید. علت امر نوعی سوداری شناختی است موسوم به لنگر انداختن.^۸

فرض کنید پژوهش نشان داده است که داروی «الف» می‌تواند علائم ۷۰ درصد مبتلایان به کرونا را به میزان ۵۰ درصد کاهش دهد. همچنین فرض کنید وقتی یافته‌ها در بیمارستان دیگری بدون اطلاع از داده‌های جمع‌آوری‌شده در بیمارستان نخست انجام شوند نشان دهند که داروی الف تنها در ۵ درصد موارد، و به میزان ۳۰ درصد، علائم بیماران مبتلا به کرونا را کاهش می‌دهد. اما در حالتی که پزشکان بیمارستان دوم از یافته‌های بیمارستان نخست مطلع باشند لنگر انداختن ذهن بر روی داده‌های اولیه سبب می‌شود که اعداد بالاتر از اعداد حالت نخست گزارش شوند. مثلاً گفته شود داروی «الف» در ۵۰ درصد مبتلایان به کرونا، و به میزان ۴۰ درصد، علائم بیماری را کاهش می‌دهد. هرچه بیمارستان نخست معروف‌تر باشد یا پزشکان دست‌اندرکار تهیه گزارش اولیه معروف‌تر باشند، لنگر انداختن نقش قوی‌تری بازی می‌کند.^۹

اما سرانجام اگر کسی درصدد آزمون دعوی چاپ‌شده برآید اولاً بخت آنکه بودجه‌ای برای پژوهش خود دریافت کند کم است و اگر موفق به کسب بودجه شود و پژوهش را مجدد تکرار کند با یک دوراهی مواجه خواهد شد: اگر یافته‌ها تکرار شوند با خوردن برچسب «کار تکراری» هرگز بختی برای چاپ نمی‌یابد و اگر یافته‌ها تکرار نشوند نیز بخت کمی برای چاپ می‌یابد؛ زیرا اذهان داورها به سمت مؤثر بودن دارو سودار شده است؛ گو اینکه ابطال آن یافته گامی به عقب است که البته سخن نادرستی است؛ زیرا با دانستن اشتباهات گامی جلورفته‌ایم.

آنکه با نادیده‌گرفتن شباهت‌های میان آزمون‌ها علمی مانند پزشکی و روان‌شناسی دیگر علمی قانون‌نهاد نخواهند بود و تبدیل می‌شوند به علمی یکه‌نگار.^۱ این درحالی‌ست که واقعاً شباهت‌های بسیار زیادی نیز بین انسان‌های موجود وجود دارد. نادیده‌گرفتن شباهت‌ها خطرات از نادیده‌گرفتن تفاوت‌هاست؛ زیرا شباهت‌های میان انسان‌ها بیش از تفاوت‌های آنهاست؛^۲ دوم آنکه در سال‌های اخیر بحران تکرارپذیری در دیگر حیطه‌ها فراتر از زیست‌شناسی و روان‌شناسی مشاهده شد که در آن تفاوت‌های فردی نیز هویدا است. به‌عنوان نمونه، مقالات حیطه‌ای مانند منابع آب و هیدرولوژی وضعیتی بسیار بدتر دارند.^۳

ج. سومین عامل تکرارپذیری را شاید بتوان به تفاوت میان آزمایشگران نسبت داد. به چنین مشکلی از دیرباز توسط فیلسوفان علم اشاره شده بود: اینکه مشاهدات خالص وجود ندارند و تفاوت در شبکه مفهومی مشاهده‌گران و پیش‌فرض‌های نظری ایشان بر آنچه مشاهده شده اثر می‌گذارد (برای مطالعه بیشتر در این مورد مقاله بودن^۴ را ببینید). اما ممکن است بخشی از مشکل در سوداری‌های شناختی همه دانشمندان جای داشته باشد. سوداری تأیید به‌عنوان یکی از معروف‌ترین سوداری‌های شناختی می‌تواند راه را بر خطاهایی نظام‌مند در علم بگشاید.^۵

ویسون در آزمایش خود که به «تکلیف انتخابی ویسون» موسوم است به نحوی ساده و با استفاده از چهار کارت بازی نشان می‌دهد چگونه در فرضیه‌آزمایی‌ها، ذهن انسان خطایی نظام‌مند انجام می‌دهد و به‌طور عمده به دنبال تأیید نظر خود است و نه آزمون درستی آن. عموماً ذهن انسان در تأیید پیش‌فرض‌های درست پنداشته شده وارد عمل می‌شود.^۶ به نظر می‌رسد که فرایند داوری مقالات راه را بر سوداری تأیید می‌بندد اما پژوهش‌ها نشان دادند که سوداری تأیید جمعی حتی شدیدتر از سوداری تأیید فردی است.^۷

1. Münsterberg, H. (1899).
2. D. J. Witherspoon et-al 2007.
3. Stagge, James H.; Rosenberg, David E.; Abdallah, Adel M.; Akbar, Hadia; Attallah, Nour A.; James, Ryan (2019-02-26).
4. Bogen, Jim (2014).
5. Oswald, Margit E.; Grosjean, Stefan (2004).
6. Wason, Peter C. (1960).
7. Lee, C.J.; Sugimoto, C.R.; Zhang, G.; Cronin, B. (2013).

8. Sherif, Muzafer; Taub, Daniel; Hovland, Carl I. (1958).

9. Jimenez, A, Mesoudi A., 2019.

بسیار نادری مبتلا هستند. کیت تشخیصی جدیدی وارد بازار شده است که یک درصد نتیجه مثبت کاذب دارد. به عبارتی از هر ۱۰۰ نفری که آزمون می‌شوند و کاملاً سالم هستند یکی را به اشتباه بیمار اعلام می‌کند. فردی مورد آزمایش قرار می‌گیرد و نتیجه آزمایش او مثبت اعلام می‌شود. احتمال بیمار بودن او چقدر است؟

عموم مردم پاسخ به این پرسش را ۹۹ درصد اعلام می‌کنند. در حالی که اگر نرخ پایه را در نظر بگیریم و در قاعدهٔ بیز قرار دهیم پاسخ عددی بسیار کوچک خواهد بود. برای آنکه نادرستی پاسخ عموم مردم مشخص شود فرض کنید صد هزار نفر این تست را داده‌اند و نتیجه تست هزار نفر مثبت بوده است. آیا باز هم خواهیم گفت که به احتمال ۹۹ درصد این هزار نفر به بیماری مبتلا هستند؟ خیر! زیرا تعداد کل مبتلایان در شهر ۱۰ نفر بیشتر نبوده است؛ اما ذهن ما چندان برای محاسبات احتمالاتی سازگاری نیافته است و نتیجه آنکه دچار خطاهای اساسی در محاسبات احتمالی می‌شود.^۷

الکساندر برد از ما درخواست می‌کند بار دیگر این پرسش را مطرح کنیم که اگر ۱۰۰ پژوهش را انتخاب کنیم که با آلفای پنج درصد چاپ شده‌اند در تکرار پژوهش‌ها باید انتظار تکرار چندتای از آنها را داشته باشیم؟ پاسخ متعارف ۹۵ مقاله نادرست است. پاسخ صحیح به نرخ پایه بستگی دارد. اگر بدانیم از هر ده مقالهٔ ارسالی به نشریات تنها یکی صادق است^۸ و این عدد را در قاعدهٔ بیز قرار دهیم پاسخ صحیح به یکباره از ۹۵ به ۶۷ مقاله کاهش می‌یابد. این خطای ادراکی ماست که انتظار داشته‌ایم یافته‌های ۹۵ مقاله تکرار شوند.

تا به اینجا برد زوایای جدیدی از بحران را معرفی کرده است. اما برد راهکاری هم دارد. به داوری او به دو شیوه می‌توان آنچه را در واقع است با آنچه انتظار داریم نزدیک کرد. نخست آنکه مقالات با کیفیت‌تری عرضه کرد که ۱۰ درصد نرخ پایه افزایش یابد؛ و دوم آنکه آلفا را از پنج‌صدم به پنج‌هزارم کاهش داد که البته، همان‌گونه که پیشتر اشاره شد، پیشنهاد دوم از برد نیست و

د. اما عموماً، و به‌ویژه در روان‌شناسی، مشکل را در روش‌های آماری می‌دانند.^۱ به این اعتبار که شاید روش‌های آماری به کار رفته در آزمون‌نظریه‌ها، روش‌های موثقی نیستند و باید مقصر اصلی را در سطح روش‌شناختی جستجو کرد. یکی از مشکلات آماری این‌گونه معرفی می‌شود که گویا بزرگ بودن ارزش پی‌رایج، احتمال خطای نوع اول را افزایش داده است و مقالاتی با داشتن آلفایی در حدود پنج‌صدم نیز امکان چاپ پیدا می‌کنند.^۲ ایشان راه‌حل را کاهش میزان ارزش پی به پنج‌هزارم پیشنهاد کرده‌اند.^۳ این پیشنهاد منتقدانی نیز داشته است که خواهان حفظ ارزش پی در همان مقدار پنج‌صدم رایج هستند.^۴ به نظر می‌رسد که راه‌حل کاهش ارزش پی نه تنها تأثیر مثبتی بر بحران تکرارناپذیری ندارد، بلکه حتی در صورت اعمال مشکل را بیشتر نیز خواهد کرد. اگر حدود پنج درصد مقالات تکرار نمی‌شدند و ما خواهان عددی کمتر بودیم این راه‌حل معقول می‌بود. اما وقتی تعداد مقالاتی که تکرار نمی‌شوند بسیار بالاتر از پنج‌صدم است نشان می‌دهد مشکل در جای دیگری وجود دارد و با تنگ‌کردن گلوگاه برای چاپ مقالات، علل دیگر، از جمله داده‌سازی که در بخش بعد از آن سخن خواهیم گفت، تشدید خواهند شد. این کار مانند سخت‌گیری بیشتر در مرزها برای جلوگیری از قاچاق کالا است بدون آنکه به وضعیت معیشتی قاچاقچیان خرد رسیدگی شود.

الکساندر برد، فیلسوف علم، به نکته آماری جالبی اشاره می‌کند که ابعاد تازه‌ای از بحران را نشان می‌دهد.^۵ وی می‌گوید که مقداری از بحران همیشه وجود داشته است و تنها متوجه آن نبوده‌ایم. به‌گمان او علت در یک سوداری شناختی انسان نهفته است موسوم به در نظر نگرفتن نرخ پایه.^۶

ابتدا با مثالی در نظر نگرفتن نرخ پایه را معرفی کنیم. جمعیت شهری ده میلیون نفر است و در شهر ۱۰ نفر به بیماری

1. Stanley, T. D.; Carter, Evan C.; Doucouliagos, Hristos (2018).
2. Denworth, Lydia (2019).
3. Benjamin, Daniel J.; et al. (2018).
4. Lakens, Daniel; et al. (March 2018).
5. Bird, Alexander. "Understanding the replication crisis as a base rate fallacy." *The British Journal for the Philosophy of Science* (2018).
6. Bar-Hillel, M., 1980. The base-rate fallacy in probability judgments. *Acta Psychologica*, 44(3), pp.211–233.

7. Kahneman, D. and Tversky, A., 1972.

8. Johnson, V. E., Payne, R. D., Wang, T., Asher, A. and Mandal, S. [2017]: pp. 1–10.

شود، بخت زیادی برای چاپ ندارند ولو ایده‌های جالبی در مقاله وجود داشته باشد یا در آن از روش‌های پژوهش بدیعی استفاده شده باشد. این سوداری به سمت نتایج را سوداری چاپ (publication bias) می‌نامند.

به‌عنوان نمونه، برخی از پژوهشگران معتقدند که فشار نهادهای عرضه‌کننده بودجه می‌تواند یافته‌های علمی را از مسیر اصلی منحرف کند.^۵ فشار دانشگاه‌ها برای چاپ هرچه بیشتر مقالات توسط پژوهشگران راه را برای برخی از سهل‌انگاری‌ها و حتی تقلب‌ها باز می‌کند. حذف داده‌های ناخواسته، داده‌سازی و دیگر شیوه‌های تقلب به‌کار گرفته می‌شوند تا مقاله‌ای باب میل نشریات به چاپ برسد و دانشگاه‌ها و مؤسسات پژوهشی در رتبه‌بندی‌ها جایگاه بالاتری پیدا کنند^۶ و بدین ترتیب بودجه‌های بیشتری را کسب کنند.

راه برون‌رفت از بحران

براساس اینکه علت اصلی بحران کدام‌یک از موارد بیان شده دانسته شود راه‌های متفاوتی به نظر می‌رسد:

الف. برای رفع مشکل مدل‌های خطی، که عوامل کمی را در بررسی یک پدیده در نظر می‌گیرند، پیشنهاد شده است که از مدل‌های وام‌گرفته از نظریه سیستم‌ها بهره گرفته شود [۱۲]. اما نکته آنجاست که قوای شناختی ما روابط چندگانه و پیچیده را به‌خوبی فهم نمی‌کند. شاید به همین دلیل است که با آنکه دیرزمانی است با نظریه سیستم‌های پیچیده آشنا شده‌ایم همچنان عمده پژوهش‌های چاپ‌شده در مجلات علمی به مقالاتی اختصاص دارند که از همان روش‌های سنتی پژوهش بهره می‌گیرند که عموماً اندرکنش دو یا سه عامل را می‌سنجد؛

ب. همان‌طور که پیش‌تر اشاره شد توجه به تفاوت‌های فردی می‌تواند راه را بر پیشرفت‌های علمی در بخش‌هایی مانند پزشکی شخصی‌شده بگشاید و احتمالاً بحران تکرارپذیری را نیز تا حدی تخفیف دهد؛

ج. وقتی از سوداری‌های نظام‌مندشناختی سخن می‌گوییم به نظر می‌رسد که از پدیده‌ای مانند وجود نقطه کور در بینایی سخن

تعداد زیادی از پژوهشگران طرفدار آن هستند. هر دو راهکار برد با مشکلاتی اساسی مواجه‌اند: نخست آنکه توصیه به عرضه مقالات با کیفیت‌تر توصیه‌ای تهی است؛ چیزی است مانند نصایح بزرگ‌تراها به کودکان که «خوب باش و بد نباش!».

البته خود برد از اصطلاح «صادق» به جای با کیفیت استفاده می‌کند که به مراتب مشکلات بیشتری دارد. هرچه نظریه‌ای محتوای قابل آزمون کمتری داشته باشد، احتمال صدق آن افزایش می‌یابد. در واقع دعوت به عرضه فرضیه‌های صادق‌تر دعوت غیرمستقیمی است به عرضه نظریات کم‌محتوا. اما پیشنهاد دوم او، یعنی کاستن از آلفا، به نحوی مستقیم تعداد مقالات چاپ‌شده را کاهش می‌دهد و بنابراین به نحوی غیرمستقیم مجدد نرخ پایه را کاهش می‌دهد و با کاهش نرخ پایه مشکل تکرارپذیری بیشتر خواهد شد. به‌علاوه مشکل بزرگتری را به همراه دارد که در بخش بعد از آن سخن خواهیم گفت.

ه. و سرانجام مهم‌تر از همه اینکه مشکل در زمینه اجتماعی علم است. می‌توان نقش عوامل اجتماعی، سیاست‌گذاری، و اقتصادی را به‌عنوان مهم‌ترین علت بحران تکرارپذیری در نظر گرفت که فشار زیادی را برای چاپ مقالات به پژوهشگران وارد می‌کند.^۱ بعد از اتمام جنگ جهانی دوم علوم به نحوی فزاینده به سمت انتشار مجلات جدید و چاپ روزافزون مقالات پیش رفت. دِرک دی‌سولا پرایس، که او را پدر علم‌سنجی می‌خوانند، در سال ۱۹۶۳ در کتاب علم کوچک، علم بزرگ^۲ پیش‌بینی کرد که با افزایش فزاینده تعداد مقالات، مقاله‌های بی‌کیفیت نیز فزونی خواهند گرفت.

تولید انبوه از توجه لازم به علم می‌کاهد. به‌ویژه وقتی مؤسسات پژوهشی و دانشگاه‌ها فشار روزافزونی بر چاپ مقالات گذاشتند اوضاع رو به وخامت بیشتری رفت. مقالات حکم کالایی را پیدا کردند که پژوهشگر با چاپ آنها امرار معاش می‌کند.^۳ از سویی دیگر، مجلات عموماً پژوهش‌ها را براساس نتایج چاپ می‌کنند؛^۴ پژوهش‌هایی که در آنها فرضیه تحقیق رد

1. Neill, U. S. (2008).

2. Price, Derek J. de Solla (1963).

3. Mirowski, P. (2011).

4. Song, F.; Parekh, S.; Hooper, L.; Loke, Y. K.; Ryder, J.; Sutton, A. J.; Hing, C.; Kwok, C. S.; Pang, C.; Harvey, I. (2010).

5. Neill, U. S. (2008).

6. Osmo Kivinen, Juha Hedman, Kalle Artukka, 2017.

ه. به نظر می‌رسد راه‌حل اساسی را باید در علت پنجم جست و فعالانه کوشید تا برخی سیاست‌گذاری‌های پژوهشی و چاپ مقالات را تغییر داد. چهار راهکار پیشنهادی را ذیل این بخش می‌توان گنجاند:

۱. بها دادن به پژوهش‌هایی که درصد بازآزمایی پژوهش‌های چاپ‌شده هستند می‌تواند ابعاد بحران را آشکارتر کند. با وجود آنکه کم‌وبیش وخامت بحران آشکار شده است، هنوز عموم مجلات معتبر علمی به دنبال نشر «پژوهش‌های دست اول» هستند و مقالاتی را چاپ نمی‌کنند که درصد تکرار نتایج چاپ‌شده بوده‌اند.^۵ در این زمینه نیاز به تغییر سیاست‌های نشر است. به‌عنوان نمونه، مجلات باید موظف شوند در بخش معرفی مجله درج کنند که مقالات مربوط به تکرارپذیری را می‌پذیرند. در حال حاضر تنها اندکی از مجلات معتبر این کار را داوطلبانه انجام داده‌اند.

افزون‌براین و مهم‌تر آنکه، مجلات موظف شوند تا درصدی از مقالات چاپ‌شده را به پژوهش‌های مربوط به تکرارپذیری اختصاص دهند. باید نگاه سیاست‌گذاران علمی به این نکته جلب شود که آگاه شدن از نادرستی پژوهش‌های گذشته کم‌ارزش‌تر از نشر داده‌های جدید نیست. انجام ندادن این کار اعتبار علم را به چالش می‌کشد و نه انجام آن، و آگاه شدن به اینکه در موارد زیادی برخطا بوده‌ایم؛

۲. سیاست بها دادن گزاف به مقالات برای ارتقاء استادان باید اساساً تغییر کند. در تولید دستوری علم، بی‌تردید، نخستین محصول تقلب است. دانشمند پرآوازه‌ای چون پیتر هیگز می‌گوید این سیاست دانشگاه‌ها باعث شد که سال‌ها نتواند کاری جدی انجام دهد و می‌گوید اگر چنین فضایی که از دهه ۱۹۹۰ در دانشگاه‌ها حاکم شده در سال ۱۹۶۴ نیز وجود داشت به سختی قابل تصور است که می‌توانست کار بر روی ذره هیگز را به پیش برد و گلايه می‌کند که کاهش توانایی او برای چاپ مقالات باعث شده است که دیگر مورد توجه دانشگاه‌ها نباشد^۶ ولو که دانشجویان می‌توانند از تجربیات او بهره گیرند.

می‌گوییم و راهی بر فائق آمدن بر آن نیست؛ درحالی‌که برخی روان‌شناسان با عرضه دلایل قانع‌کننده‌ای کارکرد مثبتی از این سوداها را در گرفتن تصمیم‌های سریع در زندگی روزمره قائلند^۱ و به همین دلیل نیز آن را قواعد هدایت ذهن می‌خوانند و نه سوداری، که باری منفی دارد، اما آشکارا به نظر می‌رسد که سوداری‌هایی مانند سوداری تأیید و در نظر نگرفتن نرخ پایه اثراتی منفی در پژوهش‌های علمی دارند که فرایندهایی متاملانه‌اند و نه سریع. آیا می‌توان بر چنین سوداری‌هایی فائق آمد؟ اگر بتوان گام مهمی است در کاستن از بحران تکرارپذیری.

روان‌شناسان برای تغییر برخی از سوداری‌های شناختی راهکارهای کارآمدی را پیشنهاد کرده‌اند.^۲ از آنجاکه این سوداری‌ها عموماً ناخودآگاه و خودکار هستند تخفیف آنها، اگر نگوییم ناممکن، کاری بس زمان‌بر است.^۳ نخستین گام برای تخفیف یا تغییر سوداری مهمی مانند سوداری تأیید، که بسیار بر روند علم تأثیر دارد، آگاهی از آن است. اما به نظر نمی‌رسد که در دروسِ روش تحقیق سرفصل‌هایی به سوداری‌های شناختی، و به‌ویژه سوداری‌هایی که بر روند کلی علم مؤثرند، وجود داشته باشد. بنابراین، ساده‌ترین و نخستین گام عملی در جهت کاستن از بحران تکرارپذیری افزودن درسی به نام آشنایی با سوداری‌های شناختی برای تمامی دانشجویان است.

د. انداختن تقصیر به گردن روش‌های آماری ساده‌ترین شیوه برای نادیده انگاشتن مقصر اصلی است. روش‌های آماری ابزارهایی هستند با خطاهای شناخته‌شده. هرچند توقع کاهش خطاها از آنها امری معقول است و از زمان پیدایی روش‌های آماری به نحوی مستمر در جهت عرضه ابزارهای نوین آماری و تصحیح ضرایب گام‌های مؤثری برداشته شده است،^۴ اما وقتی خطاهایی بسیار فراتر از میزان پیش‌بینی شده در یک ابزار ریاضاتی مشاهده می‌کنیم آشکارا مشکل را باید در جایی دیگر جست. باوجوداین، سیاست‌گذاران کلان در دانشگاه‌ها و مؤسسات پژوهشی ترجیح می‌دهند تا تقصیر را به گردن آمار و آماردانان بیندازند.

1. Gigerenzer, G. (2006).

2. Beard, C. (2011).

3. Ariely, D. (2008).

4. Stigler, Stephen M. (1986).

5. Martin, G. N.; Clarke, Richard M. (2017).

6. Aitkenhead, Decca. Peter Higgs: I wouldn't be productive enough for today's academic system, The Guardian 6 December 2013.

درمی‌آیند. برداشتن شرط چاپ مقاله برای دفاع از رساله دکتری دانشجویان دکتری می‌تواند گام بسیار مهمی برای جلوگیری از چاپ مقالاتی توسط ایشان باشد که به احتمال زیاد در تکرارپذیری نتایج خوبی نخواهند گرفت؛

۳. کاستن از رقابت‌های علمی میان کشورها و افزودن بر همکاری‌های بین‌المللی می‌تواند از شدت بحران بکاهد. طی دهه‌های گذشته درباره نقش مثبت و منفی رقابت در علم نظرات متعددی بیان شده است (برخی از این نظرات را می‌توانید در مقاله فنگ و کیسِدوال ببینید).^۲ هرچند رقابت می‌تواند در مواردی نتایج مثبتی برای پیشرفت علم داشته باشد، اما بحران تکرارپذیری نشان داد فضای رقابتی بسیار شدیدی که بر علم امروز حاکم است عموماً نتایج منفی داشته است. از راه‌های کاستن از این فضای رقابتی شدید و افزودن بر همکاری می‌توان به جنبش «علم باز» اشاره کرد. جالب آنکه ایده انتشار مجله‌های علمی با ایده علم باز آغاز شد؛ اما طی سال‌های اخیر که با بحران تکرارپذیری مواجه شده‌ایم، برخی راه برون‌رفت از مشکل را شفاف‌سازی در علم دانسته‌اند.^۳

وقتی پژوهشگر می‌داند نتایج واقعی پژوهش وی، باوجود آنکه فرض تحقیق وی قبول یا رد شده باشد، چاپ می‌شود و داده‌های خام نیز در اختیار عموم قرار می‌گیرد، دیگر نه فشاری برای رسیدن به نتایج «مطلوب» احساس می‌کند و نه دستکاری در داده‌هایی انجام می‌دهد که به نظاره عموم می‌رسد؛ اما علم باز تنها در صورتی موفق خواهد بود که تغییراتی پیشنهادی در قوانین بخش قبل انجام شود.

بحران کرونا نشان داد که همکاری‌های بین‌المللی و در اختیار قرار دادن اطلاعات خام امری کاملاً ضروری است. اگر چینی‌ها توالی کشف‌شده ژنوم کرونا و اویروس را مطابق سیاست علم باز در اختیار عموم قرار نمی‌دادند، هر کدام از شرکت‌های تولیدکننده واکسن مجبور بودند به نحوی مستقل به کشف این توالی بپردازند و این ورود نخستین واکسن کرونا را با تأخیر بیشتری روبه‌رو می‌کرد. بنابراین علم باز نه تنها مشکل بحران تکرارپذیری را تخفیف می‌دهد، بلکه به پیشرفت سالم علم نیز کمک می‌کند [۴۷].

چنین فضایی راه را برای چاپ مقالات به شیوه‌های ناسالم باز می‌کند. وقتی دانشمندان نام‌آور علم، مانند هیگز، اسیر چنین سیاست‌هایی باشند کاملاً مشخص است که پژوهشگران جوان نیز کاملاً در سیطره چنین جو سنگینی مجبور به چاپ مقالات به هر شیوه ممکنه شوند تا امنیت مالی داشته باشند. اینجاست که «هارکینگ» سالم‌ترین شیوه انحراف از پژوهشی سالم است.

نوربرت کر، روان‌شناس اجتماعی، در سال ۱۹۹۸ مقاله‌ای دارد^۱ که در آن اصطلاح هارکینگ (HARKing: hypothesizing after the results are known) را معرفی می‌کند: اینکه پژوهشگر پس از آنکه داده‌ها را جمع کرده است تازه شروع به نگارش فرضیه می‌کند. چنین فرضیه‌ای در واقع به قامت داده‌ها دوخته شده و مشخص است که ابطال نمی‌شود و در نهایت چاپ می‌شود؛ اما در واقع چنین فرضیه‌ای به نحو مستقل آزمون نشده است و چه بسا اگر بخواهیم مستقل آن را آزمایش کنیم احتیاج به مدل آزمایشی کاملاً متفاوتی باشد. اگر فشار نهادهای سفارش‌دهنده مقاله نباشد، پژوهشگر دست به چنین اقداماتی نخواهد زد.

پژوهشگران جوان برای استخدام در پژوهشگاه‌ها و دانشگاه‌ها نیازمند رزومه‌ای با ده‌ها مقاله چاپ‌شده هستند. پس از استخدام نیز برای حفظ شغل و ارتقاء در آن باید هر سال با چاپ چند مقاله رزومه را «پربارتر» کنند. معیارهای رقابت در بازار آزاد به تدریج جای معیارهای متعارف پژوهش علمی را گرفته‌اند. در سایه بحران تکرارپذیری، قوانین استخدامی کنونی برای پژوهشگران و اعضای هیئت علمی، و بخشنامه‌های ارتقاء و ترفیع استادان نیازمند بازنگری اساسی‌اند.

در ایران از شرط‌های پذیرش موضوع برای پایان‌نامه‌های کارشناسی ارشد و رساله‌های دکتری این است که موضوع تکراری نباشد. علت مخالفت‌ها برای انجام تکرار یک پژوهش منتشرشده آن است که در تکرار یک پژوهش همان نتایج قبلی به دست خواهد آمد؛ اما طی دو دهه اخیر این پیش‌فرض به چالش کشیده شده است. زمان آن است که پیش‌شرط تکراری نبودن پایان‌نامه‌ها و رساله‌ها بازبینی شود.

افزون‌براین، درصد قابل توجهی از مقالات نشریافته توسط ایرانی‌ها مقالاتی هستند که توسط دانشجویان دکتری به نگارش

2. Fang FC, Casadevall A. 2015.

3. Collaboration, Open Science (2012).

1. Kerr, N. L. (1998).

می‌یابند و اصلاح آن خطاها امری بس دشوار است. اصلاح برخی دیگر، از جمله روش‌های آماری، همواره نزد خود آماردانان در حال انجام بوده است و خواهد بود و به‌علاوه توصیه به اینکه آماردان‌ها باید روش‌های آماری بهتری تولید کنند بی‌وجه است. اما مشکل پنجم در معنای عام‌تری از روش‌شناسی مطرح می‌شود؛ یعنی به معنایی که فیلسوفان و جامعه‌شناسان علم از روش‌شناسی در نظر دارند. از آنجاکه شرایط اجتماعی تولید مقالات علمی با تغییر آیین‌نامه‌ها و دستورالعمل‌ها قابل تغییر است، بنابراین عملی‌ترین راه برای مقابله با بحران تکرارپذیری بازنگری در سیاست‌های چاپ مقالات و نقش چاپ مقالات در ارتقاء جایگاه علمی پژوهشگران است.

در پایان اشاره‌ای به کیفیت پایین داوری مقالات نیز اهمیت دارد. حجم بسیار بالای تعداد مقالات باعث شده است تعداد زیادی داوری به داوران داده شود و این باعث می‌شود کیفیت داوری‌ها به دلیل خستگی داوران افت کند. نام این پدیده را خستگی داور (reviewer fatigue) گذاشته‌اند. در نتیجه مجلات مجبورند از تعداد بیشتری از داوران دعوت کنند. در گزارشی که در مجله نیچر به چاپ رسیده^۱ نشان می‌دهد طی سال‌های ۲۰۱۳ تا ۲۰۱۷ تعداد دعوت از داوران جدید بسیار بالا رفته و هم‌زمان کیفیت داوری‌ها و نرخ پذیرش داوری توسط داوران کاهش داشته است. بنابراین، حجم بالای مقالات باعث افت کیفیت داوری‌ها شده و در نتیجه می‌تواند از علل تشدید بحران تکرارپذیری باشد.

نتیجه‌گیری

در این مقاله به بحرانی اشاره شد که علم با آن دست‌به‌گریبان است. بخشی از علل را باید در روش‌شناسی به معنای محدودتر کلمه، یعنی آن‌گونه‌که در دروس روش تحقیق در دانشگاه‌ها تدریس می‌شود، جست. چهار علت روش‌شناختی در این معنای محدود معرفی شدند. نخستین علت اینکه تفاوت در شرایط آزمایش و به‌ویژه در تفاوت در مدل‌های آزمایش می‌تواند از علل تکرارناپذیری باشد. گفته شد دومین علت را باید در اختلاف آزمودنی‌های موجود در گروه‌های آزمایشی پی‌گرفت. سومین علت را در تفاوت میان آزمایشگران معرفی کردیم و سرانجام دیدیم باید چهارمین علت را در روش‌های آماری مورد استفاده پی‌گرفت. آگاهی یافتن از این چهار علت روش‌شناختی تا حدی می‌تواند از شدت بحران بکاهد.

اما به دو دلیل علت پنجم مهم‌تر تشخیص داده شد. علت پنجم این بود که بحران را عموماً باید در زمینه اجتماعی تولید مقالات علمی جست. دلیل نخست آنکه تمامی چهار علت نخست همواره با علم جدید وجود داشته‌اند؛ اما به بحرانی که امروزه شاهد آن هستیم نینجامیده‌اند؛ دلیل دوم آنکه برخی از آن علل، از جمله سوداری‌های شناختی، به سرشت انسان ربط

1. Peer reviewers unmasked: largest global survey reveals trends" 7 Sep 2017

منابع

12. Evidence-Based Medicine Working Group (1992), "Evidence-based medicine. A new approach to teaching the practice of medicine", *JAMA*, 268 (17): 2420–25.
13. Fang FC, Casadevall A. (2015), "Competitive science: is competition ruining science? *Infect Immun*", 83: 1229–1233.
14. Feyerabend, P.K., (1959), "An Attempt at a Realistic Interpretation of Experience", in P.K. Feyerabend, *Realism, Rationalism, and Scientific Method* (Philosophical Papers I), Cambridge: Cambridge University Press, 1985, pp. 17–36.
15. Geoffrey S. Ginsburg, Huntington Ph.D Willard (2009), *Essentials of Genomic and Personalized Medicine*, Academic Press.
16. Gigerenzer, G. (2006), "Bounded and Rational", *Contemporary Debates in Cognitive Science*, R. J. Stainton, Blackwell Publishing: 115–133.
17. Hanson, N.R., (1958), *Patterns of Discovery*, Cambridge, Cambridge University Press.
18. Ioannidis, John P. A. (August 1, 2005), "Why Most Published Research Findings Are False", *PLOS Medicine*. 2 (8): e124.
19. Jimenez, A, Mesoudi A., (2019), "Prestige-biased social learning: current evidence and outstanding questions", *Palgrave Communications*, 5(1).
20. Johnson, V. E., Payne, R. D., Wang, T., Asher, A. and Mandal, S. (2017), "On the Reproducibility of Psychological Science", *Journal of the American Statistical Association*, 112, pp. 1–10.
21. Kahneman, D. and Tversky, A., (1972), "Subjective probability: A judgment of representativeness", *Cognitive psychology*, 3(3), pp.430–454.
1. Aitkenhead, Decca (2013), "Peter Higgs: I wouldn't be productive enough for today's academic system", *The Guardian*, 6 December.
2. Ariely, D. (2008), *Predictably Irrational: The Hidden Forces That Shape Our Decisions*, Harper Collins.
3. Bar-Hillel, M., (1980), "The base-rate fallacy in probability judgments", *Acta Psychologica*, 44(3), pp. 211–233.
4. Beard, C. (2011), "Cognitive bias modification for anxiety: current evidence and future directions", *Expert Review of Neurotherapeutics*, 11(2), 299–311.
5. Benjamin, Daniel J.; et al. (2018), "Redefine statistical significance", *Nature Human Behaviour*. 2 (1): 6–10.
6. Bird, alexander (2020), "understanding the replication crisis as a base rate fallacy", *the british journal for the philosophy of science*.
7. Boyd, Nora Mills and James Bogen, "Theory and Observation in Science", *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/fall2021/entries/science-theory-observation/>
8. Collaboration, Open Science (2012), "An Open, Large-Scale, Collaborative Effort to Estimate the Reproducibility of Psychological Science", *Perspectives on Psychological Science*, 7 (6): 657–660.
9. Collins, H., (1985), *Changing Order: Replication and Induction in Scientific Practice*, London: Sage Publications.
10. D. J. Witherspoon et-al (2007), "Genetic Similarities Within and Between Human Populations", *Genetics* 176(1): 351–359.
11. Denworth, Lydia (2019), "A Significant Problem: Standard scientific methods are under fire. Will anything change? ", *Scientific American*, Vol. 321, No. 4, pp. 62–67.

35. Osmo Kivinen, Juha Hedman, Kalle Artukka, (2017), "Scientific publishing and global university rankings. How well are top publishing universities recognized?", *Scientometrics*, Vol.112, pp. 679-695.
36. Oswald, Margit E.; Grosjean, Stefan (2004), "Confirmation bias", in Pohl, Rüdiger F. (ed.), *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory*, Hove, UK: Psychology Press, pp. 79-96.
37. "Peer reviewers unmasked: largest global survey reveals trends", (7 Sep 2017), *Nature News available at <https://www.nature.com/articles/d41586-018-06602-y>*.
38. Pickering, A. (1986, January), Against correspondence: A constructivist view of experiment and the real. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* (Vol. 1986, No. 2, pp. 196-206). Philosophy of Science Association.
39. Price, Derek J. de Solla (1963), *Little science, big science*, New York: Columbia University Press.
40. Rouse, Joseph (1998), "Heideggerian philosophy of science", *Routledge Encyclopedia of Philosophy*, Taylor and Francis.
41. Sherif, Muzafer; Taub, Daniel; Hovland, Carl I. (1958), "Assimilation and contrast effects of anchoring stimuli on judgments", *Journal of Experimental Psychology*, 55 (2): 150-155.
42. Song, F.; Parekh, S.; Hooper, L.; Loke, Y. K.; Ryder, J.; Sutton, A. J.; Hing, C.; Kwok, C. S.; Pang, C.; Harvey, I. (2010), "Dissemination and publication of research findings: An updated review of related biases", *Health Technology Assessment (Winchester, England)*. 14 (8): iii, iix-xi, iix-193.
43. Stagge, James H.; Rosenberg, David E.; Abdallah, Adel M.; Akbar, Hadia; Attallah, Nour A.; James, Ryan (2019-02-26), "Assessing data availability and research reproducibility in hydrology and water resources", *Scientific Data*, 6: 190030.
22. Kerr, N. L. (1998), "HARKing: Hypothesizing after the results are known", *Personality and Social Psychology Review*, 2 (3): 196-217.
23. Kuhn, T.S. (1962), *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press, reprinted 1996.
24. Lakens, Daniel; et al. (March 2018), "Justify your alpha", *Nature Human Behaviour*, 2 (3): 168-171.
25. Lee, C.J.; Sugimoto, C.R.; Zhang, G.; Cronin, B. (2013), "Bias in peer review", *Journal of the Association for Information Science and Technology*, 64: 2-17
26. Lorenz, Edward N. (March 1963), "Deterministic Nonperiodic Flow", *Journal of the Atmospheric Sciences*. 20 (2): 130-141.
27. Martin, G. N.; Clarke, Richard M. (2017), "Are Psychology Journals Anti-replication? A Snapshot of Editorial Practices", *Frontiers in Psychology*, 8: 523.
28. Mirowski, P. (2011), *Science-Mart: Privatizing American Science*, Harvard University Press.
29. Münsterberg, H. (1899), "Psychology and history", *Psychological Review*, 6, 1-31.
30. Nature Video (28 May 2016), "Is There a Reproducibility Crisis in Science?", *Scientific American*, <https://www.scientificamerican.com/video/is-there-a-reproducibility-crisis-in-science/>.
31. Neill, U. S. (2008), "Publish or perish, but at what cost?", *Journal of Clinical Investigation*, 118 (7): 2368.
32. Neill, U. S. (2008), "Publish or perish, but at what cost?", *Journal of Clinical Investigation*. 118 (7): 2368.
33. Orrell, David (2012), *Truth or Beauty: Science and the Quest for Order*, New Haven: Yale University Press.
34. Orrell, David; Smith, Leonard; Barkmeijer, Jan; Palmer, Tim (2001), "Model error in weather forecasting", *Nonlinear Processes in Geophysics*, 9 (6): 357-371.

- and Brain, and Its Implication for Questions of Generalization and Replication”, *Minds and Machines*, 28 (2): 353–374.
47. Wason, Peter C. (1960), “On the failure to eliminate hypotheses in a conceptual task”, *Quarterly Journal of Experimental Psychology*, 12 (3): 129–140.
48. Weir, Kristen (2011), “A reproducibility crisis?”, *American Psychological Association*.
44. Stanley, T. D.; Carter, Evan C.; Doucouliagos, Hristos (2018), “What meta-analyses reveal about the replicability of psychological research”, *Psychological Bulletin*, 144 (12): 1325–1346.
45. Stigler, Stephen M. (1986), *The History of Statistics: The Measurement of Uncertainty Before 1900*, Harvard University Press.
46. Wallot, Sebastian; Kelty-Stephen, Damian G. (2018), “Interaction-Dominant Causation in Mind

