

Abstract

What is 'research quality'? Some answers

In May 2023, experts from different scientific disciplines came together at Berlin-Brandenburg Academy of Sciences and Humanities to discuss perspectives on quality in research



The replication crisis sparks intense debate: Aiming to ensure quality in research, many initiatives have been launched, the CoARA process being one of the most recent. However, it is the concept of quality itself that comes under scrutiny: Is it expert opinions that ensure quality in research, or adhering to certain standards? Should quality be measured by excellence or impact? And is there such a thing as a concept of quality applicable to all of science?

Five experts from different academic fields met at Berlin-Brandenburg Academy of Sciences and Humanities to debate what quality in research entails and what can be done to promote it. The panel was chaired by Ulrich Dirnagl (BIH).

In her opening remarks, **Anna Dreber** stressed that replicability may be interpreted in a number of ways. For her, replication means testing established hypotheses using similar methods with new data. However, in economics, this is often very difficult, leading to more focus on reproducibility where the same data is studied as in the original study. “There is no other way as we cannot, e.g., go back in time, divide Germany into East and West Germany, reunite it again and then see whether we get the same results.” This makes it hard to evaluate the reliability of results in terms of replication success, Dreber explained. Reliability and credibility of results, however, could still be assessed by conducting multiverse analyses, sharing data and altering the publication process. “I think so called Registered Reports are an important step to generate high quality publications.”

Health researcher **Gordon Guyatt** reflected on the relationship of science and truth, highlighting the historical evolution of concepts and ideas in science. “For a long time, we have been convinced that randomized trials show us which interventions are effective. Now, this has become controversial as some claim – although very misguidedly – that looking at

large databases may bring us closer to the truth.” He also noted that guidelines for clinicians used to be written by expert groups without attention to standards that we now consider crucial to creating trustworthy guidelines: ensuring diverse panels, stating questions precisely, and searching for evidence in a systematic fashion. The lack of a gold standard made it impossible to know which approach might be correct, the researcher argued. “We have current guidelines that meet our standards trustworthy, but it’s hard to conceptualize that as getting closer to the truth.” Guyatt concluded that quality in research might be impossible to define yet recognized when seen.

“Aristotle described quality as a fluid concept that could mean different things in different contexts”

According to the sociologist **Lena Hipp**, the very straightforward answer to the question of what high quality research is, relates to common scientific standards such as methodological soundness, novelty, creativity, and relevant, perhaps even surprising results. In light of an ongoing replication crisis, the thin slicing of research results to publish many, though closely related papers, and the emergence of predatory journals, however, she argued that the quality of research in the social sciences should not be defined by its findings. Moreover, she suggested to put more emphasis on pre-registering reports, to strengthen gatekeeping mechanisms, and to reduce pressures on researchers to publish as many papers as possible. “This creates false incentives that potentially lead to poor research.”

Christoph Markschies began by pointing out that the humanities have difficulties with a concept of quality standards that is determined primarily or even exclusively by the category of replicability. However, he also emphasized that even Aristotle described quality as a fluid concept that could mean different things in different contexts. In this respect, he said, it is not surprising if certain elements of quality are relevant to all scientific disciplines, while other elements are important only to a few. Outstanding achievements in the humanities (and cultural studies) are always characterized by novelty (novelty in the sense of originality), not by replication or replicability. “Replication is obviously much more important for natural and technical sciences because of the importance of data collection and data interpretation.” Nevertheless, Markschies warned against banishing the concept of replication/replicability altogether from discussions of quality standards in the humanities, thereby undermining the guiding idea of a kind of unity of all sciences in certain regards. Certain results of a source criticism in the historical-philological disciplines must of course be replicable in order to develop criteria for better and worse interpretations. For this reason alone, it is necessary to be precise in dealing with quality problems in the humanities and cultural studies as well.

“Quality in research is defined by originality, relevance, and technical soundness”

Susanne Schreiber argued that quality in research is defined by originality, relevance, and technical soundness. Whether a paper is judged as original and relevant, however, largely depended on personal opinion. “It boils down to the decisions of editors and often enough the standing of the scientists involved.” To promote impartiality, the neuroscientist favoured double blind reviews and saw some potential in open publishing. To promote technical soundness, Schreiber proposed allocating special funds that enable scientists to make their data more accessible to colleagues. She also argued for putting greater emphasis on the

reusability of data. “There are complex, expensive studies in which, e.g., brain connectivity is deciphered by using electron microscopes. Appropriate resources are needed to make this data truly accessible and reusable for other groups.”

In the ensuing debate, chair Dirnagl wondered whether representatives of all scientific disciplines should put their heads together to compile different standards of quality in a list. Markschieß welcomed the idea, meanwhile Dreber saw more potential in generating such lists in between disciplines working with similar methods. Furthermore, she advertised for establishing a practice of Registered Reports to eliminate publication bias: In this model, researchers submit their proposals like extended pre-analysis plans before conducting their research, and journals agree to publish the studies no matter the actual later results. “This is the way to make sure we’re testing interesting hypotheses in credible ways and avoid the file-drawer problems by having more null results published.”

In response to Schreiber’s opening statement, Guyatt proposed that only technical soundness should be a criterion of quality in the natural sciences as an excessive focus on originality, e.g., was discouraging researchers from conducting replication studies of great importance and value. “People should not be penalized for doing replications which are essential to ultimately arriving at the truth, particularly for estimates of treatment effect.” Markschieß remarked that, in the humanities, originality was a useful, high standard element of quality and therefore could and should not be eliminated. “To remove originality would produce a term of quality which, I’m sorry to say, is useless for the humanities.”

Wrapping up the debate, Schreiber questioned whether originality was solely relevant within the humanities or rather an important element of quality within the natural sciences as well. “It’s when something original is done that society moves forward. Just having studies that do not produce new insights, yet are technically sound, can be bad for science.” She went on to question whether large amounts of publications adhering to standards would produce the desired outcomes. “Publishing according to high scientific standards and usability is important, no question, but producing ever more publications without innovative content for me isn’t a solution.” The neuroscientist concluded her argument by advertising for reducing the number of papers published each year meanwhile increasing the quality of publications.