BJHS Themes (2026), 1–19
doi:10.1017/bjt.2026.10041

RESEARCH ARTICLE

The regulatory ethos in science

Nicole C. Nelson[1] and Lara Keuck[2]

[1]Department of Medical History and Bioethics, University of Wisconsin–Madison, USA and [2]Institute for Studies of Science, Department of History, Philosophy and Theology, and Medical School OWL, Bielefeld University, Germany
Corresponding author: Nicole Nelson; Email: nicole.nelson@wisc.edu

Abstract

This paper introduces the concept of the *regulatory ethos* to describe some common values and ideals that underscore the close connection between validation and regulation that this issue of *BJHS Themes* explores. We identify the primary motivation for this ethos as making knowledge production processes traceable, and define the regulatory ethos as valuing plans over situated actions, uniformity over heterogeneity, auditing over communication, and validation over validity. Standard operating procedures, reporting checklists, preregistrations, compliance rules and monitoring are key means through which this ethos is enacted. While regulators have been instrumental in promulgating this ethos, it is not confined to the regulatory sphere. We argue that reforms aimed at enhancing rigour and reproducibility are an example of how the practices and values associated with regulatory science have diffused out into academic science. Identifying this ethos as regulatory in origin – rather than wholly new or as part of a broader process of modernization – allows us to see that alternatives are not unscientific per se, and better identify the strengths and weaknesses of the regulatory ethos of science, such as the risk that data produced through these procedures will be replicable, statistically rigorous, and transparent, but not meaningful.

In 2021, the science foundation of Berlin inaugurated an annual prize of €500,000 together with the Center for Responsible Research of the Berlin Institute of Health (BIH) and several additional funders and institutions. The Einstein Foundation Award for Promoting Quality in Research is awarded to an individual, an institution, and an early-career project.[1] In their own words, the award 'aims to provide recognition and publicity for outstanding efforts that enhance the rigor, reliability, robustness, and transparency of research in the life sciences, natural sciences, the social sciences, and the humanities, and stimulate awareness and activities fostering research quality among scientists, institutions, funders, and politicians.'[2]

[1] The awarded individuals in 2021 were Paul Ginsparg, the founder of the preprint server arXiv; in 2022 Gordon Guyatt, a pioneer of evidence-based medicine; in 2023 Yves Moreau, a big-data engineer advocating for ethical standards of gathering and using human DNA data; and in 2024 Elisabeth Bik, who identifies image manipulations in biomedical papers. The awarded institutions in 2021 were the Center for Open Science, in 2022 the Psychological Science Accelerator, in 2023 the Berkeley Initiative for Transparency in the Social Sciences and in 2024 the PubPeer. The awarded early-career projects in 2021 were ManyBabies, in 2022 Ape Research Index, in 2023 the Responsible Research Assessment Initiative and in 2024 PixelQuality.

[2] Einstein Foundation, 'About', at https://award.einsteinfoundation.de/about (accessed 10 January 2025).

The Einstein Foundation Award for Promoting Quality in Research is financially half the size of a Nobel Prize, but is very different in its aims and scope. It is not dedicated to the discovery of a groundbreaking research result or a revolutionary scientific method, but rewards people and institutions who have reformed the ways in which normal research is pursued, evaluated, and published. The award is driven by a scientific social movement of quantitative metascience studies and science activists who are concerned about the integrity, reproducibility and quality of research.[3] The concern and appreciation of scientific reforms is testimony to steep career of what we call the *regulatory ethos* in science. This paper describes and historicizes this development with regard to both the advent of pursuing and evaluating science according to norms typical of *regulatory* contexts, and the framing of these norms in form of an *ethos*.

The papers in this issue of *Themes on Testing Knowledge: Validation and Regulation in the Health and Human Sciences* provide many examples of how scientific methods have been couched in terms of quality control, standard procedures and regulatory certification. Our goal in this paper is to describe the specific ethos of science that these examples illustrate, how they contrast with other visions of how to secure the credibility and authority of science, and how this ethos is being taken up in new scientific arenas. In what follows, we first motivate why we chose to work with the term 'regulatory ethos', and describe how regulations and regulatory agencies have been pivotal in coalescing a specific set of norms and practices for the conduct of science. Second, we characterize this ethos with respect to four central features of the concept, namely its preference for planability, uniformity, accountability and reliability. Finally, we turn to the case study of contemporary rigour and reproducibility reforms to show how the regulatory ethos can exist in areas of science with very different structures of oversight which are distant from policymaking processes. Indeed, one of our central motivations for describing the regulatory ethos is so that we can more clearly see instances in which it has taken root outside regulatory science, and for some scientific reformers (such as the architects of the Einstein Foundation Award) has become synonymous with the ethos of science itself. We conclude with what we see as one of the main dangers of this trend – that the regulatory ethos promotes the pursuit of validation sometimes even at the expense of validity, thereby potentially producing results that are replicable, statistically rigorous and transparent, but not meaningful.

Introducing 'ethos' and the role of regulatory science

The term 'ethos' has a long history in philosophy, and is usually characterized by drawing on Aristotle's definition of *êthos*, which denotes a person's credibility as a function of his character. In science and technology studies (STS) and history and philosophy of science (HPS), the term has gained traction through the work of the sociologist Robert Merton (after whom was named the Robert K. Merton Center for Science Studies, a Berlin-based institution that has close ties to the Einstein Foundation Award). According to Merton, 'ethos' means a set of norms that contain 'moral as well as technical prescriptions' for how scientists should behave.[4] They are moral in that they give guidance about what is good, and they are technical in that they are 'procedurally effective'. The Mertonian ethos is thus a set of qualities and values that are meant to secure the status, credibility and authority of the people and institutions who stand for this ethos. According to Judy Segal and Alan Richardson, Merton self-consciously appropriated Aristotle's notion of ethos to apply it to

[3] David Peterson and Aaron Panofsky, 'Metascience as a scientific social movement', *Minerva* (2023) 61(2), pp. 147–74.

[4] Robert K. Merton, 'The normative structure of science', in Norbert W. Storer (ed.), *The Sociology of Science: Theoretical and Empirical Investigations*, Chicago: University of Chicago Press, 1973, pp. 267–78.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

the scientific *persona*. For them, the crucial question that a so-understood scientific ethos replies to is, 'why should we believe science?'[5] We agree that this understanding of ethos is a more helpful framing of Merton's concept than treating it as a particularly adequate sociological approach to describing academic professionalization.

A critical insight from the literature on objectivity is that there is no single answer to Merton's question of why we should believe science. Rather, as Lorraine Daston and Peter Galison have argued, there are distinct forms of objectivity which are each responsive to specific fears about what might make science untrustworthy.[6] While Merton's norms are attracting renewed attention in discussions of scientific reform, Bart Penders and colleagues have critiqued scientific reformers' use of Mertonian norms in particular and virtue-theoretic frameworks in general as oversimplifications which do not recognize that there are different ways of enacting a virtue or that there are different moral programmes travelling together under the banner of reform.[7]

STS and HPS scholars have instead offered analyses that give specificity to current scientific reforms, either by placing them in the context of broader historical trends or by emphasizing their continuity with previous styles of objectivity. Jill Morawksi, for example, has argued that the move towards 'institutionalized regulative solutions' to failures to replicate in psychology represents an extension of mechanical objectivity.[8] For her, the key continuity is the fear of subjectivity and the desire to eradicate the various cognitive biases that plague psychological science. Just as Arthur Waddington (a central actor in Daston and Galison's description of mechanical objectivity) was distressed to find that his beliefs about the symmetry of nature had influenced his perception of splashing droplets, so too, she argues, are contemporary reformers distressed to see evidence of how confirmation bias or hindsight bias have influenced their data analysis practices. While Morawski sees reformers' use of regulatory and auditing practices as part and parcel of mechanical objectivity, Penders treats contemporary reform movements as part of a broader history of scientific bureaucratization.[9] For Penders, the adoption of bureaucratic practices is not simply about reducing subjectivity, but distinguishing oneself and one's science as being more advanced and civilized than other forms of science. The concept of an ethos speaks to status and authority as well as to epistemic concerns.

Daston and Galison's description of mechanical objectivity shares many commonalities with our description of the regulatory ethos, such as the prescription that scientists should follow 'a set of procedures that would ... move nature to the page through a strict protocol'.[10] But, in contrast to Morawski, we argue that the fear that drives the regulatory ethos is not that scientists will impose their 'unchecked will onto nature'; it is that they will do so in a way that is not *traceable*. As we will show, human-induced bias is not a problem per se as long as one can measure and document the bias. The fear is failing to report properly, committing fraud by not cohering to the preregistered plan, failing the audit.

[5] Judy Segal and Alan W. Richardson, 'Introduction. Scientific ethos: authority, authorship, and trust in the sciences', *Configurations* (2003) 11(2), pp. 137–44.

[6] Lorraine Daston and Peter Galison, *Objectivity*, New York: Zone Books, 2007.

[7] Mohammad Hosseini, Enric Senabre Hidalgo, Serge P.J.M. Horbach, Stephan Güttinger and Bart Penders, 'Messing with Merton: the intersection between open science practices and Mertonian values', *Accountability in Research* (2024) 31(5), pp. 428–55; Mare Knibbe, Sarah de Rijcke and Bart Penders, 'Care for the soul of science: equity and virtue in reform and reformation', *Cultures of Science* (2025) 8(1), pp. 12–23.

[8] Jill Morawski, 'Psychologists' psychologies of psychologists in a time of crisis', *History of Psychology* (2020) 23, pp. 176–98, 189.

[9] Bart Penders, 'Process and bureaucracy: scientific reform as civilisation', *Bulletin of Science, Technology & Society* (2022) 42(4), pp. 107–16.

[10] Daston and Galison, op. cit. (6), pp. 120–1.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

Other scholars have already emphasized discontinuity with older scientific norms when studying regulatory science practices. Alberto Cambrosio and colleagues, for example, introduced the term 'regulatory objectivity' to describe what they argue is a distinctive set of norms.[11] To them, the key distinction is that regulatory objectivity operates at the level of the collective – it is an ethos embodied not by an individual practitioner, but by a group. Cambrosio and colleagues' description of regulatory objectivity is connected to our concept of the regulatory ethos, particularly in its focus on the development of procedures to coordinate collective action and the pursuit of compatibility and coordination over the pursuit of truth. Jeremy Freese and David Peterson describe what they call 'statistical objectivity' similarly, although they point to a shift from individual studies to groups of studies (rather than from individuals to groups of people) as their key feature.[12] We argue, however, that the regulatory ethos has spread beyond the regulatory agencies, standard-setting organizations and consensus conferences that Cambrosio and colleagues studied, and has become an ideal to be embodied by individuals as well as groups. Cambrosio and colleagues argue further that collective emphasis of regulatory objectivity means that it is less concerned with the relationship between the observer and nature, and more concerned with coordinating action. As they put it, 'under the regime of regulatory objectivity, biomedicine suspends the search for "true values" and replaces it with the establishment of conventions … it is less important to arrive at a truth (analytic or otherwise) than it is to ensure compatibility between different laboratories and different hospitals'.[13] This second point begins to depart from the very idea of a search for a mind-independent reality. This can be regarded as a further expansion of the concept of objectivity (as Cambrosio and colleagues argue), or can raise the question whether the norms and practices they describe should still be analysed under the rubric of objectivity. Our concept of regulatory ethos can be used as an analytic category to describe the phenomenon independently of this question.

The regulatory ethos sits at the intersection of regulatory objectivity, mechanical objectivity, bureaucratization and a Mertonian ideal of good science, making it at times difficult to see how the regulatory ethos is not just an enactment of one of these ideals. We argue that this is not a mere analytic difficulty for STS/HPS, but that the confusion of these overlapping norms and processes has in fact had a productive role in allowing the regulatory ethos to migrate outside regulatory science. From many different vantage points, the regulatory ethos can simply be mistaken as the *scientific* ethos, not a specific form of arguing for the trustworthiness of science to which there are viable alternatives.

A good deal of scholarship already exists in STS which attempts to define regulatory science based on the context in which science is conducted or its orientation towards policy, to highlight the wide range of activities which could plausibly count as regulatory science, and to contrast regulatory science with other types of science.[14] Our aim is not to contribute to this literature. Rather, we treat sites of regulatory action as separable from the ethos itself so that we can better understand how the moral and technical prescriptions of the

[11] Alberto Cambrosio, Peter Keating, Thomas Schlich and George Weisz, 'Regulatory objectivity and the generation and management of evidence in medicine', *Social Science & Medicine* (2006) 63(1), pp. 189–99; Alberto Cambrosio, Peter Keating, Thomas Schlich and George Weisz, 'Biomedical conventions and regulatory objectivity: a few introductory remarks', *Social Studies of Science* (2009) 39(5), pp. 651–64.

[12] Jeremy Freese and David Peterson, 'The emergence of statistical objectivity: changing ideas of epistemic vice and virtue in science', *Sociological Theory* (2018) 36(3), pp. 289–313.

[13] Cambrosio *et al.*, op. cit. (11), p. 195.

[14] Alan Irwin, Henry Rothstein, Steven Yearley and Elaine McCarthy, 'Regulatory science: towards a sociological framework', *Futures* (1997) 29(1), pp. 17–31; Sheila Jasanoff, *The Fifth Branch: Science Advisers as Policymakers*, Cambridge, MA: Harvard University Press, 1998; Liora Salter, 'Introduction to mandated science', in Salter (ed.), *Mandated Science: Science and Scientists in the Making of Standards*, Dordrecht: Springer Netherlands, 1988, pp. 1–19.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

regulatory ethos are increasingly not limited to regulatory science contexts, but extend to the evaluation of research and researchers in general.

Defining this ethos as regulatory is important because it allows us to identify links between regulatory science and contemporary reform movements in academic science, which in turn allows us to use the existing literature on regulatory science to better understand these developments. In the United States, agencies such as the Food and Drug Administration (FDA), the Environmental Protection Agency (EPA) and the US Department of Agriculture (USDA) have been key sites for the creation of regulatory standards for conducting scientific studies and analysing data, such as the Clinical Laboratory Improvement Amendments standard (which regulates laboratories performing clinical tests on human samples) and the Good Laboratory Practice standards (which regulates the production of data on the safety and/or efficacy of chemicals such as pesticides or pharmaceutical drugs).[15] The twentieth century has seen worldwide growth in agencies and committees dedicated to the quality assurance of products and processes. Work conducted by the World Health Organization's International Agency for Research on Cancer, for example, intersected with discussions at the US EPA on whether (and, if so, which) Petri dish or animal test is sufficiently valid and feasible for qualifying that a chemical can cause cancer and therefore should be legally regulated.[16] Localizing the promulgation of the specific form of this ethos of science within the institutionalization of regulatory bodies can help to address the history, sociology and politics of the moral project and its epistemological implications.

Regulatory agencies have been successful in defining and disseminating a specific ethos in no small part because of their role as gatekeepers to the market – they can compel would-be participants to collect, analyse and disseminate data in a particular way. The presence of systems to monitor compliance has further reinforced regulators' preferred mode of conducting science. Simply put, regulatory agencies have the institutional power to make others adopt their ethos. This does not mean that this ethos is confined to activities that regulatory agencies are legally empowered to regulate. As Daniel Carpenter has argued in his history of the FDA, the agency's gatekeeper role has meant that 'its mere suggestions and intimations induce compliance' from drug developers, resulting in the widespread adoption of practices even in the absence of legal enforcement.[17] The diffusion of regulatory standards also takes place within institutions such as an individual company, where following standards such as Good Laboratory Practice standards can streamline operations within the company and provide documentation that a company adhered to the rules, thereby reducing the company's liability. Indeed, as Jeremy Greene has shown, the establishment of a distinct 'FDA science' contributed to developing clinical trials into an industry for the newly founded contract research organizations in the late twentieth century.[18]

Many instances of regulatory science may only partly embody the ethos that we describe, and our concept of the regulatory ethos is not meant to be synonymous with

[15] For a history of Good Laboratory Practice see Colleen Lanier-Christensen, 'Between science and administration: regulatory epistemology, corporate influence, and endocrine-disrupting chemicals', *DePaul Law Review* (2025) 74(2), pp. 399–460.

[16] Thomas, this issue; Creager, this issue. See also Valentin Thomas, *Classé cancérogène: Enquête sur un processus entravé*, Paris: Presses de Sciences Po, 2025.

[17] Daniel P. Carpenter, *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*, Princeton, NJ: Princeton University Press, 2010. On the emergence of regulatory agencies' mode of 'guidance' instead of setting formal rules see also Kevin M. Lewis, 'Informal guidance and the FDA', *Food and Drug Law Institute* (2011) 66(4), pp. 507–50. For a broader history on forms and understandings of formal and informal rules see Lorraine Daston, *Rules: A Short History of What We Live By*, Princeton, NJ: Princeton University Press, 2022.

[18] Jeremy A. Greene, *Prescribing by Numbers: Drugs and the Definition of Disease*, Baltimore: Johns Hopkins University Press, 2007.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

the actual actions of regulatory agencies, or of the scientists who produce data for regulatory purposes. This, we believe, is one of the strengths of the concept: creating some analytical separation between regulatory science and the practices and ideals commonly associated with this work allows us to more easily identify similarities with work in differently regulated sectors, and to consider how these ideals and practices might function in spaces without the same market incentives, legally authorized gatekeepers or compliance-monitoring systems. Taking this perspective can aid us in considering how, for example, the practice of preregistration functions differently in research regulated by the FDA compared to research regulated by institutional ethics committees, and can help us to elucidate which elements of a regulatory ethos are replaced and which might still be enforced in contemporary political appeals to so-called deregulation. The motivation of our characterization of the elements of this ethos, and our discussion of their enactment in various case studies in the twentieth and twentieth-first centuries, is to open up a broader debate about the origins of this ethos, its historical and contemporary forms and enactments, and the implications of and alternatives to this kind of thinking about good science.

The four features of the regulatory ethos: planability, uniformity, accountability, reliability

The first feature of the regulatory ethos is that, to borrow a phrase from Lucy Suchman, it favours plans over situated actions.[19] In this view, good scientists should specify in advance the steps that they will take (as in the case of a clinical trial) or follow a pre-specified series of steps (as in the case of test guidelines or standard operating procedures). Once specified, these plans take on rule-like qualities that govern scientists' conduct, and the execution of the plan becomes an end in itself. In contrast to a situated model of action, where the trajectory of science is expected to evolve as scientists engage with the world, the unexpected moments that inevitably arise as scientists attempt to carry out their plans are seen as problems to troubleshoot.[20] Rather than privileging iteration, improvisation or a more expansive notion of problem solving that involves finding new paths to get to a goal, the regulatory ethos encourages scientists to bring their actions back in line with the original plan to the extent possible and to document any deviations.

An example from medical science that illustrates the virtue of following plans is provided by Jingwen Li's paper in this issue, which reconstructs a lawsuit brought by a customer against his optician in 1920s Singapore.[21] Unhappy with the glasses he had been prescribed, the customer sued, claiming that the prescription was faulty. As evidence that the doctor had behaved appropriately, the defence showed that he had followed proper procedures, analysing the patient's eyes with a retinoscope and prescribing accordingly. The plaintiff's lawyer, in contrast, argued that what distinguished a correct from an incorrect prescription was not the procedure followed, but how the patient felt when wearing the resulting glasses. A competent optician should conduct a retinoscopy examination but should also modify the calculated prescription as needed to produce a suitable pair of glasses for that patient. What this case and its interpretations show is that a focus on following rules tends to privilege specific kinds of evidence, namely evidence that is produced by mechanical, interrater reliable procedures, as well as a distinct kind of expert who knows how to follow the plan correctly. If the plan nonetheless does not work out, the responsibility can

[19] Lucy A. Suchman, *Plans and Situated Actions: The Problem of Human–Machine Communication*, Cambridge and New York: Cambridge University Press, 1987.

[20] Andrew Pickering, *The Mangle of Practice: Time, Agency, and Science*, Chicago: University of Chicago Press, 1995; Hans-Jörg Rheinberger, *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*, Stanford, CA: Stanford University Press, 1997.

[21] Li, this volume.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

be delegated to external factors, such as the unruly customer. This feature of the regulatory ethos aligns closely with mechanical objectivity, in that it idealizes practitioners who behave in machine-like ways rather than those who exercise their trained judgement. Yet Li's case study reminds us also of the limits of objectifying the subjective experience of vision correction.

A second feature is that the regulatory ethos favours uniformity over heterogeneity, difference and locality. Uniformity encompasses two kinds of prioritization: first, it privileges a broad accessibility of methods that does not depend heavily on individual performance and adjusting of technical and epistemic resources. Second, uniformity prioritizes framings of the objects of analysis that abstract away their idiosyncrasies, and focuses on objectified parameters. Uniformity goes hand in glove with strict adherence to plans, since alignment with pre-specified plans is more difficult to achieve when the conditions of work differ. In this view, the ideal plan is one that can be executed in a variety of locations and does not require access to highly specialized reagents, technologies or workspaces. Likewise, the ideal scientist is one who is interchangeable for any other similarly trained scientist, not one who has a particular aptitude for executing a technique or who has cultivated a unique skill set over the years. In this view, the world should be treated as (or remade into) a uniform terrain where any reasonably trained person can carry out a protocol and get the same result. The regulatory ethos has a democratizing impulse that changes the character of both science and scientists: expertise becomes a matter of discipline rather than skill, and specialized techniques become a liability rather than an advantage because they are heterogeneously distributed.

Colleen Lanier-Christensen's study illustrates this ethos at work in the selection of methods of the Association of Official Agricultural Chemists (AOAC) for assessing the chemical composition of fertilizer.[22] In considering potential methods to adopt as a standard, the AOAC wanted methods that were accurate, precise and reliable, but also 'appropriate for routine use'. Lanier-Christensen shows that when those two goals were in conflict, the AOAC prioritized uniformity. The United States Commissioner of Agriculture agreed, arguing that it would be preferable to use a 'less perfect' method than a more precise method that only some laboratories would be equipped to perform. Importantly, Lanier-Christensen stresses the difference between quality assurance as being only about the process, and quality control, which is directed at the outcome. She understands the introduction of Good Laboratory Practice standards in the 1970s as a form of quality assurance without quality control that has prioritized toxicological studies which follow these – contested but standardized – procedures over investigative studies that have tested alternative hypotheses but don't have the same status in regulatory and juridical contexts as the (often industry-friendly) procedural studies do.[23]

Angela Creager's study of work by federal agencies to evaluate potential laboratory tests for chemical carcinogenicity likewise shows a preference for accessible methods.[24] A National Cancer Institute committee charged with evaluating tests indicated that they would not support a patented test when alternatives were available in the public domain. The fact that the inventor of the eponymous Ames test had begun sharing his bacterial strains even before publishing on them was a factor in its eventual adoption as a regulatory standard, since it meant that government, industry and academic scientists alike had access to the materials needed to carry out the protocol. Historical studies of the technique used to standardize the measurement of electrical resistance reveal a similar pattern. The British used metal coils as a means of measurement because they were well suited to use by

[22] Lanier-Christensen, this volume.
[23] Lanier-Christensen, op. cit. (15), p. 399.
[24] Creager, this volume.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

telegraphists, even though metal coils were not as accurate as other available techniques.[25] The German mercury-based technique offered greater precision but at the cost of radically decreased accessibility, and when the German Reich legalized the ohm in 1898, they chose a less precise but more widely reproducible version of the technique as the basis for their standard.[26] In contrast to the examples discussed in outlining the first feature, the actors in all of these cases resist the pursuit of ever more refined means of producing mechanical impressions of nature. These examples align more strongly with regulatory objectivity, since they sacrifice potential precision for the ability to coordinate action across different domains.

A third feature of the regulatory ethos is that it treats documenting and reporting as processes that enable auditing rather than communication. Conceptualized as a communicative tool, a methods section is an imperfect means of transmitting information, and the shortcomings of this tool can be repaired through mechanisms such as face-to-face communication. Conceptualized as an audit trail, a methods section is a chronological record of actions that allows a third party to detect deviations from a protocol, problematic uses of methods, or even instances of fraud. Incomplete methods information is not a communicative gap to be repaired but a warning sign that may cast doubt on the credibility of the findings. This element of the regulatory ethos reflects the logic of accounting that has not only formed modern economics but also transformed evaluations writ large, including the evaluation of education, science and health care.[27] Accounting pairs well with the aforementioned feature of planned action: in accounting, the yardstick of right and wrong is compliance to the initial plan. This is a very different register of trustworthiness than building trust on trained judgement, interpersonal relationships, practical experience or tacit knowledge, all of which challenge the exchangeability and uniformity of labourers that the regulatory ethos prioritizes. Reporting, documenting and compliance have a double role of delimiting sources of bias and creating a written trail of evidence that testifies that the research process has been done correctly.

Christopher Phillips's paper exemplifies this aspect by showing how the FDA propagated off-the-shelf statistical tools for documenting that a rational and uniform procedure had been performed to determine the effects of a chemical.[28] This use of statistical tools allowed regulators to define a so-called 'virtual safe dose' of a chemical despite the inherent uncertainties of such measurements. It did not eliminate the problems of identifying a practical zero residue or a lowest detectable amount, but it did allow regulators to move forward by documenting the assumptions that they were making and the steps they were taking in light of known uncertainties. The goal of such documentation was not to communicate to other scientists how they might perform similar assessments, but to transparently disclose how the FDA had acted in light of existing variability and uncertainty. This example illustrates the distinction we drew earlier between the fear of bias itself (which drives mechanical objectivity) and the fear that bias will not be properly disclosed, which is what we see as central to the regulatory ethos.

[25] Simon Schaffer, 'Late Victorian metrology and its instrumentation: a manufactory of Ohms', in Robert Bud and Susan Cozzens (eds.), *Invisible Connections: Instruments, Institutions and Science*, London: SPIE Press, 1992, pp. 23–56.

[26] Kathryn M. Olesko, 'Precision, tolerance, and consensus: local cultures in German and British resistance standards', in Jed Z. Buchwald (ed.), *Scientific Credibility and Technical Standards in 19th and Early 20th Century Germany and Britain*, vol. 1, Dordrecht: Springer Netherlands, 1996, pp. 117–156.

[27] Axel C. Hüntelmann and Oliver Falk (eds.), *Accounting for Health: Calculation, Paperwork, and Medicine, 1500–2000*, Manchester: Manchester University Press, 2021; Cris Shore and Susan Wright, *Audit Culture: How Indicators and Rankings Are Reshaping the World*, London: Pluto Press, 2024; Marilyn Strathern, *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*, New York: Routledge, 2000.

[28] Phillips, this volume.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

This leads us to the final feature: the belief that if every step of the research process has been done correctly – planning action, reproducible conduct of the plan, careful documenting and reporting – this will result in validated results *and therefore* valid knowledge. The fact that the process is validated *vouches* for the validity of its results (the vouching expression stems from Nancy Cartwright, who characterized a similar epistemic expectation with respect to the privileged role of randomized controlled trials to generate extrapolatable results[29]). In the most extreme case, the validity of a research result (i.e. that the result is actually informative of what it is intended to be about) becomes secondary to the validation of the process (i.e. that the result is reproducible).

Yingying Han's examination of the rise and fall of the forced swim test is an example of how the validity of this behavioural test became secondary to its broad use as a reliable validation tool that provided comparable data for assessing antidepressants.[30] Animal activists, however, successfully referred to scientific studies that questioned the validity of the relationship between an important symptom of human depression and the time rodents spent finding a platform in a water pool before giving up to make the case against the use of this test in depression research. According to them, the forced swim test led to invalid extrapolation from animals to humans, and therefore should lose its status as a standard test, and even be banned. In a similar vein, Creager argues that while scientists and regulators were able to validate the Ames test as a method, they were less successful in their efforts to validate it as a model.[31] They succeeded in developing standardized protocols that generated reproducible results across laboratories, but identified several limitations of the test as a tool for identifying compounds that might cause cancer in humans. A chemical could cause cancer without causing mutations, for example, which meant that the Ames test missed known carcinogens such as asbestos. Phillips's actors likewise objected to the claim that the 'virtual safe dose' was synonymous with actual safety, or that the calculation could be straightforwardly mobilized in larger regimes of risk calculation to protect public health.

In Creager and Phillips's cases, we see how tests became embedded in the regulatory apparatus without fully erasing the question marks around their validity. Lanier-Christensen's case, in contrast, offers a clear example where actors prioritized validation over validity. Industry and government actors both recognized that a reliable, reproducible test could serve an important market coordination function even if that test was not the most accurate measure of the actual content of fertilizer. As long as each fertilizer was measured by the same yardstick, consumers could still make choices between products and companies could benefit from fair, predictable assessments of their products. These examples demonstrate two pathways through which validation can supplant validity – in the first case, because the practical necessity of producing science for regulation allowed validation to carry the day even in the absence of validity, while in the second case validity was deprivileged even though it was available because validation was more aligned with the needs and interests of the actors within the sphere being regulated. Han's case demonstrates that this alignment is a reversible process and needs to be continuously maintained and hedged: with the diversification of interests and actors, the scientific legitimacy of a decade-long validation consensus between regulatory agencies, scientific communities and pharmaceutical companies came under pressure, and the never-resolved debate about the validity of the forced swim test gained new momentum.

[29] Nancy Cartwright, *Hunting Causes and Using Them: Approaches in Philosophy and Economics*, Cambridge: Cambridge University Press, 2007.

[30] Han, this volume.

[31] Creager, this volume.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

Taken together, these four features and their illustrative cases highlight an important function of the regulatory ethos: it serves as a compass that guides knowledge production and decision making by striving to create trans-spatial and transtemporal processes that will generate generalizable facts. As Ted Porter and others have shown, early instances of regulation were important for building consumer trust in products that circulated beyond the environs in which they were produced.[32] Standardized bushels of grain or protocols for measuring antibodies allowed for commercial trade, public-works projects, medical practice and more to move forward in an increasingly democratized and globalized world. Standards, rules, bureaucratic processes and systems of measure worked because, as Porter put it, they 'extract[ed] a severe discipline from [their] users' and thus had a 'powerful appeal to the wider public' because they impl[ied] 'personal restraint'.[33] They are critical to establishing what Geoffrey Hosking calls 'strong thin trust' – that is to say, trust in institutions, of which we often know little, which define rules and enforce rule following.[34] And the public display of that rule following serves as a technology of distinction that marks out good from bad, advanced from primitive, and civilized from uncivilized – both in science and elsewhere.[35]

Historians might question how new this development is and point to older means of establishing authoritative norms for science and technology through quasi-regulatory bodies such as the academies of sciences. Surely there are undercurrents to the advancement of the regulatory ethos that have a much longer history than the one that is connected to the rise of regulatory agencies in the past century, which we focus on in this paper. However, as Melinda Baldwin has shown with respect to the prehistory of peer reviewing – and Merton's (and Harriet Zuckerman's) misunderstanding of its epistemic role in academies' periodicals of the seventeenth century – the same practice can be attuned to quite different ideals: while 'refereeing' originally meant receiving the advice of other scientists or scholars, the renaming of this practice as 'peer reviewing' in the 1970s reflected a shift towards gatekeeping and rules of procedure to assess the *quality* of scientific research.[36] The long history, but also changing understanding, of the role and ideal of reviewing is thus not a counterexample but rather an additional illumination of the rise of the regulatory ethos of science. It is no coincidence that this ethos spread with the institutionalization of national or international bodies that manage and regulate science, such as national funding agencies (cf. Baldwin's example of peer reviewing), or regulatory agencies that act as gatekeepers to market access and certifiers for the security or efficacy of chemicals and biologicals.

In the next section, we examine the recent history of reproducibility reforms to demonstrate how the slippage between the regulatory ethos and the scientific ethos writ large has facilitated the spread of norms and practices commonly associated with regulatory science into other sectors. Many of the features that we have defined as sitting in opposition to regulatory science – evolving plans, results that depend on highly specialized skills or techniques, reliance on interpersonal trust rather than documented rule following – are features that have come under heavy criticism in the last decade. We show how reformers have reached for techniques common in regulatory science, such as preregistration and reporting checklists, as a means of cultivating a regulatory ethos outside regulatory science.

[32] Theodore M. Porter, *Trust in Numbers*, reprint edn, Princeton, NJ: Princeton University Press, 1996.

[33] Porter, op. cit. (32), p. 5.

[34] Geoffrey Hosking, *Trust: A History*, Oxford: Oxford University Press, 2014.

[35] Penders, op. cit. (9).

[36] Melinda Baldwin, 'Peer review', *Encyclopedia of the History of Science* (2020) 4(1), at https://ethos.lps.library.cmu.edu/article/id/19 (accessed 1 March 2026).

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

The replication crisis and the regulatory ethos

The Dutch psychologist Eric-Jan Wagenmakers described 2011 as a 'year of horrors' for psychology.[37] First, the *Journal of Personality and Social Psychology* (*JPSP*) – a top journal for the field – published an article by Daryl Bem claiming that people can see into the future. Second, a researcher whom Wagenmakers described as one of social psychology's 'brightest young stars', Diederick Stapel, was found to have been fabricating data for decades.[38] Both of these events, in Wagenmakers's view, pointed towards problems with how psychologists conducted, analysed and peer-reviewed experiments. If standard processes failed to detect dozens of fraudulent papers but accepted the existence of extrasensory perception, what did that say about the integrity of those processes? 'One may well wonder about the scientific status of a field in which an academic serial-killer can wander loose, undetected, for decades,' Wagenmakers mournfully concluded.[39]

Wagenmakers was particularly galled that *JPSP* had refused to publish manuscripts that failed to replicate Bem's results, arguing that knowledge of failed replications would have aided the field in detecting problems in both Bem's and Stapel's work, since 'from presented work it is often completely unclear to what degree the data were tortured to obtain the reported confession'.[40] While Wagenmakers focused on replications as a tool to identify problematic scientific results, others focused on disclosure itself as the core problem to be solved. If it was evident from the beginning exactly how much a data set had been 'tortured', then problematic findings could be identified even without conducting a replication.

Joseph Simmons, Leif Nelson and Uri Simonsohn took this second approach, focusing on documenting and making visible existing practices rather than changing them.[41] They designed an experiment to show that standard methods of statistical analysis could prove an impossible hypothesis: that research participants were younger if they listened to the Beatles than if they listed to a control song. The trick was that the participant's age was only one of many dependent variables that they measured: participants' father's age, mother's age, gender, political orientation, self-reports of how much they would enjoy eating at a diner and several other variables were also measured but showed no differences between the Beatles intervention and control groups. In a particularly impactful figure, they presented what they called a 'requirement-compliant' abstract that included all the variables in grey text with the selected information that would be included in a typical abstract in bold. If researchers were required to report all the greyed-out information rather than just the variables that produced statistically significant associations, it would be immediately obvious to readers that the one association was likely a chance variation rather than a true finding.

Simmons and colleagues' interpretation of the forces driving this selective reporting of information is in line with Morawski's analysis of the 'psychologizing' framing of what would come to be known as the replication crisis.[42] They argued that most psychologists were not wilfully attempting to deceive their readers; they were simply self-serving in their interpretations, and pointed to a large body of research from psychology and elsewhere showing that people are 'remarkably adept at reaching justifiable conclusions that mesh

[37] Eric-Jan Wagenmakers, 'A year of horrors', *De Psychonoom* (2012) 27, pp. 12–13.

[38] Wagenmakers, op. cit. (37), p. 12.

[39] Wagenmakers, op. cit. (37), p. 12.

[40] Wagenmakers, op. cit. (37), p. 13.

[41] Joseph P. Simmons, Leif D. Nelson and Uri Simonsohn, 'False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant', *Psychological Science* (2011) 22(11), pp. 1359–66.

[42] Morawski, op. cit. (8).

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

with their desires'.[43] Simmons and colleagues' solutions, however, did not call for reining in these biases, but merely for documenting their presence. They proposed reporting requirements as a solution that left 'the right and responsibility of identifying the most appropriate way to conduct research in the hands of researchers, requiring only that authors provide appropriately transparent descriptions of their methods so that reviewers and readers can make informed decisions regarding the credibility of their findings'.[44] In a retrospective piece published several years after their original article, they noted that several journals had adopted their reporting requirements, although they noted with amusement one that had rejected their suggestions because it threatened to 'dull … some of the joy scholars may find in their craft'.[45]

Reporting requirements are a widely used tool of regulators, and a tool that embodies many aspects of the regulatory ethos. FDA requirements for reporting are vast, and include, for example, reports on clinical trials testing investigational new drugs, adverse events or deaths from drugs or devices, or how much of a drug a manufacturer has produced or packaged. These requirements homogenize what might otherwise be unstructured and heterogeneous information. For example, because a chemical compound can be described by a structural formula, molecular formula, generic name, trade name and so forth, regulated companies are required to provide a CAS registry number that uniquely identifies around 300 million chemical substances. Reporting requirements are also built to enable auditing rather than communication, providing the FDA with information needed to manage risk and monitor the drug supply chain.

Reporting checklists were one of the first tools that reformers reached for in the early days of the replication crisis, and Simmons and colleagues were far from the only ones to make this suggestion. As Nelson and colleagues have shown, while written discussions of replication and reproducibility emphasize different causes and consequences of the problem, they tend to agree on transparency and open-science measures as the solution.[46] An early conference at the US National Institutes of Health on rigour and reproducibility resulted in the publication of paper that the authors framed as a 'call for transparent reporting', which specified four features that should be reported in all biomedical papers.[47] Nature Publishing Group followed suit in an editorial policy change announced in 2013, which eliminated space limits on methods section and introduced a reporting-checklist requirement for all authors.[48] In their editorial announcing this change, they noted that while not all scientists had access to facilities to validate their reagents or the resources to increase their sample sizes, they did have the capacity to produce 'a full report of how a study was designed, conducted and analysed that will allow reviewers and readers to adequately interpret and build on the results'.[49]

[43] Simmons, Nelson and Simonsohn, op. cit. (41), p. 1360.

[44] Simmons, Nelson and Simonsohn, op. cit. (41), p. 1363.

[45] Joseph P. Simmons, Leif D. Nelson and Uri Simonsohn, 'False-positive citations', *Perspectives on Psychological Science* (2018) 13(2), pp. 255–9, 259.

[46] Nicole C. Nelson, Kelsey Ichikawa, Julie Chung and Momin M. Malik, 'Mapping the discursive dimensions of the reproducibility crisis: a mixed methods analysis', *PLOS ONE* (2021) 16(7), Public Library of Science, p. e0254090, at https://doi.org/10.1371/journal.pone.0254090.

[47] Story C. Landis, Susan G. Amara, Khusru Asadullah, Chris P. Austin, Robi Blumenstein, Eileen W. Bradley, Ronald G. Crystal, Robert B. Darnell, Robert J. Ferrante, Howard Fillit, Robert Finkelstein, Marc Fisher, Howard E. Gendelman, Robert M. Golub, John L. Goudreau, Robert A. Gross, Amelie K. Gubitz, Sharon E. Hesterlee, David W. Howells, John Huguenard, Katrina Kelner, Walter Koroshetz, Dimitri Krainc, Stanley E. Lazic, Michael S. Levine, Malcolm R. Macleod, John M. McCall, Richard T. Moxley III, Kalyani Narasimhan, Linda J. Noble, Steve Perrin, John D. Porter, Oswald Steward, Ellis Unger, Ursula Utz and Shai D. Silberberg, 'A call for transparent reporting to optimize the predictive value of preclinical research', *Nature* (2012) 490(7419), pp. 187–91.

[48] 'Announcement: reducing our irreproducibility', *Nature News* (2013) 496(7446), p. 398.

[49] 'Announcement: reducing our irreproducibility', op. cit. (48).

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

Preregistration is another favoured tool of replication reformers, as Penders has explored, and like reporting requirements it is both used by regulators and embodies the regulatory ethos.[50] Preregistration takes reporting a step further by requiring researchers to state in advance how they will collect data, analyse it and judge the success of their experiment. Deviations from the preregistered protocol are to be avoided where possible and documented where unavoidable, and the presence of the preregistration – held by an independent third party and available for scrutiny by others – vouches for the integrity of the experiment and the knowledge that it has produced.

Until recently, formal preregistration practices were largely confined to the context of clinical trials, and regulatory agencies and quasi-regulatory actors played an important role in promulgating the practice. A widely acknowledged inflection point in the history of preregistration practices is a 2005 decision by the International Committee of Medical Journal Editors (ICMJE) to require preregistration in a public database as a condition of publication in member journals.[51] Since the ICMJE's membership included very high-profile journals such as the *New England Journal of Medicine* and the *Journal of the American Medical Association*, this policy had an immediate effect: the largest registry of public trials at the time, clinicaltrials.gov, had registered a total of 13,153 clinical trials from its inception in 2000 until May 2005; in the five months following the ICMJE's announcement this total jumped 73 per cent to a total of 22,714.[52] The ICMJE acted because of what they perceived as a failure of regulation: the lack of a system for tracking clinical trials meant that pharmaceutical companies could simply choose not to publish the results of trials that did not cast their products in a favourable light. *The Lancet* lambasted the FDA for overseeing a process that was 'so easily manipulated by those with potentially massive financial gains'.[53] The ICMJE's quasi-regulatory mandate was superseded in the United States only a few years later by the FDA Amendments Act of 2007, which made trial registration compulsory for a more expansive set of clinical trial types and required the disclosure of trial results. The existence of the clinicaltrials.gov infrastructure itself is also a product of regulation: the 1997 Food and Drug Administration Modernization Act of 1997 mandated the creation of a publicly accessible database of clinical trials that were linked to an investigational new drug application to the FDA.

This is not to say that the regulatory ethos cannot spread in the absence of regulation, and psychology is a good case in point. The American social psychologist Brian Nosek has been at the forefront of what he and colleagues have called the 'preregistration revolution'.[54]

In an early position paper, they pointed to preregistration in drug development as a practice to emulate, and argued that the model by which most science operates would be considered 'laughably quaint for ensuring responsibility and accountability in state or corporate governance'.[55] Under Nosek's leadership, the Center for Open Science has pursued a strategy for change that is focused on building infrastructure and community to 'make

[50] Penders, op. cit. (9).

[51] Catherine D. DeAngelis, Jeffrey M. Drazen, Frank A. Frizelle, Charlotte Haug, John Hoey, Richard Horton, Sheldon Kotzin, Christine Laine, Ana Marusic, A. John P.M. Overbeke, Torben V. Schroeder, Hal C. Sox, Martin B. van der Weyden, and International Committee of Medical Journal Editors, 'Clinical trial registration: a statement from the International Committee of Medical Journal Editors', *Journal of the American Medical Association* (2004) 292(11), pp. 1363–4.

[52] Deborah A. Zarin, Tony Tse and Nicholas C. Ide, 'Trial registration at ClinicalTrials.gov between May and October 2005', *New England Journal of Medicine* (2005) 353(26), pp. 2779–87.

[53] 'Depressing research', *The Lancet* (2004) 363(9418), p. 1335.

[54] Brian A. Nosek, Charles R. Ebersole, Alexander C. DeHaven and David T. Mellor, 'The preregistration revolution', Proceedings of the National Academy of Sciences (2018) 115(11), pp. 2600–6.

[55] Nosek *et al.*, op. cit. (54), p. 625.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

it easy' and 'make it normative' for researchers to engage in preregistration and other behaviours they view as desirable.[56] Penders describes this strategy as one that is based in appeals to the moral character of the scientist, as evidenced by public displays of emotions such as shame and embarrassment about having failed to use such practices in the past.[57] And, arguably, this strategy has worked. While preregistration was 'almost unheard of' in psychology in 2012, a recent study of psychology articles estimated that 14 per cent of articles published in prominent journals now contain preregistrations.[58]

As we have argued above, one advantage of identifying the desired ethos of replication reformers as regulatory is that it allows us to leverage the existing scholarship on regulatory science to better understand how the practices associated with the regulatory ethos may function differently when placed in new institutional contexts. For example, the history of regulatory science suggests that change can be fragile in the absence of legal mandates. Creager's study of the International Committee for Protection against Environmental Mutagens and Carcinogens (ICPEMC) in this issue is a case in point. The scientists who participated in the committee recognized the precarity of their work, fearing that the extensive (and expensive) data that they produced might be lost in time. As Creager puts it, they 'recognized that certifying testing regimes was not sufficient for regulation'. This self-organized commission of scientists did not have the authority to negotiate legal change – or the responsibility to implement policy.[59] And in the end, after more than a decade of work developing and validating screening tests for identifying cancer-causing chemicals, the committee folded in the 1990s due to lack of funding.

Reformers are not unaware of these risks. While Nosek and colleagues have long held that policy interventions to make behaviours 'required' should follow from earlier transformational work rather than acting as the catalyst for transformation, they have also long cautioned against reforms that rely solely on ideals. In their paper on the preregistration revolution, they noted that most areas of science have relatively weak incentives for adopting new practices, in contrast to the legal requirements for clinical trial registration, and called on funders and journals to exercise the same kind of quasi-regulatory power as the ICMJE to encourage the adoption of preregistration.[60]

Nosek's own experiences with preregistration reveal the limits of transplanting the regulatory ethos outside its infrastructural ecosystem. In September 2024, *Nature Human Behaviour* retracted a paper because, although the authors claimed that the analyses in the paper were preregistered, a closer inspection of the preregistration itself revealed that the study had originally been designed to test a different hypothesis.[61] Ironically, the unregistered titular claim was that practices such as preregistration could increase the replicability of the social-behavioural sciences, and the paper's authors included Brian Nosek and Charles Ebersole, both of whom were also authors on the aforementioned paper on the preregistration revolution. In a commentary about the retraction, the authors argued that the

[56] Brian Nosek, 'Strategy for culture change', at www.cos.io/blog/strategy-for-culture-change (accessed 24 September 2024).

[57] Penders, op. cit. (9).

[58] Uri Simonsohn, '[115] Preregistration prevalence', at http://datacolada.org/115 (accessed 20 November 2024); Tom E. Hardwicke, Robert T. Thibault, Beth Clarke, Nicholas Moodie, Sophia Crüwell, Sarah R. Schiavone, Sarah A. Handcock, Khanh An Nghiem, Fallon Mody, Tuomas Eerola and Simine Vazire, 'Prevalence of transparent research practices in psychology: a cross-sectional study of empirical articles published in 2022', *Advances in Methods and Practices in Psychological Science* (2024) 7(4), at doi:10.1177/25152459241283477.

[59] Creager, this volume.

[60] Nosek *et al.*, op. cit. (54).

[61] John Protzko, Jon Krosnick, Leif Nelson, Brian A. Nosek, Jordan Axt, Matt Berent, Nicholas Buttrick, Matthew DeBell, Charles R. Ebersole, Sebastian Lundmark, Bo MacInnis, Michael O'Donnell, Hannah Perfecto, James E. Pustejovsky, Scott S. Roeder, Jan Walleczek and Jonathan W. Schooler, 'Retraction note: high replicability of newly discovered social-behavioural findings is achievable', *Nature Human Behaviour* (2024) 8(10), p. 2067.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

statement that all analyses had been preregistered was indeed erroneous, and that the article should be corrected to reflect that only some of the analyses had been preregistered.[62] But others saw the situation as much more serious because the unregistered analyses were key to the central claim of the paper. 'It is one thing to forget to preregister', one commentator wrote, 'but it is entirely different to simply switch outcomes and pretend you were studying that thing all along'.[63] The editors of *Nature Human Behaviour* agreed, citing 'lack of preregistration for measures and analyses supporting the titular claim' and 'selection of outcome measures and analyses with knowledge of the data' as reasons for retracting the article.[64]

It is not just the absence of legal enforcement which means that even the most ardent proponents of the regulatory ethos will sometimes fail to consistently embody it. Preregistration gains its power in large part because of the presence of regulators who are employed to analyse the data made available through this process and empowered to act on their findings – a labour force which is lacking outside the context of clinical trials. On its website, the FDA describes the 2007 Amendments Act which mandated trial registration as an act that would 'ensure that FDA staff have the additional resources needed to conduct the complex and comprehensive reviews necessary to new drugs and devices'.[65] In contrast, much of the spadework for the investigation of the *Nature Human Behavior* article was conducted by academics acting in a personal capacity. Joe Bak-Coleman, one of the scholars who was most deeply involved in the investigation, wrote a blog post titled 'I did not want to write this blog post', where he detailed the difficulties of doing this work outside his regular employment:

> Mostly, I don't have the time to write this. I work a 9-to-5, am writing a book, and have science to do. My website broke so I've slapped it up temporarily on a placeholder website ... I don't have time to fix it. I suspect the editors, my co-authors, reviewers, and ethics board members who spent countless hours trying to understand what the authors did and correct the record felt much the same.[66]

As scholars studying equity in open sciences practices have noted, the availability of data does not translate directly into benefit: potential users still need skills, time and other resources to extract value from the data, and some are better equipped with these resources than others.[67] Likewise, as scholars studying reproducibility reforms have noted, the availability of a preregistered protocol has a very different value when the review process is managed by a largely voluntary workforce of academics than it does in the presence of a dedicated workforce of civil servants.[68] This incomplete migration of the infrastructure that allows the regulatory ethos to work can ironically stabilize the enactment of this ethos in

[62] John Protzko, Brian Nosek, Sebastian Lundmark, James E. Pustejovsky, Nicholas Buttrick and Jordan Axt, 'Historical summary of "High replicability of newly discovered social-behavioral findings is achieveable" paper', Open Science Framework (25 September 2024), at https://osf.io/2s94g (accessed 2 March 2026).

[63] Joseph Bak-Coleman, 'I did not want to write this blog post', at https://joebakcoleman.com/blog/2024/protzko (accessed 2 March 2026).

[64] Protzko *et al.*, op. cit. (61).

[65] US Food and Drug Administration, 'Food and Drug Administration Amendments Act (FDAAA) of 2007', *FDA*, at www.fda.gov/regulatory-information/selected-amendments-fdc-act/food-and-drug-administration-amendments-act-fdaaa-2007 (accessed 10 January 2025).

[66] Bak-Coleman, op. cit. (63).

[67] Ciara Staunton, Carlos Andrés Barragán, Stefano Canali, Calvin Ho, Sabina Leonelli, Matthew Mayernik, Barbara Prainsack and Ambroise Wonkham, 'Open science, data sharing and solidarity: who benefits?', *History and Philosophy of the Life Sciences* (2021) 43(4), article 115.

[68] Penders, op. cit. (9).

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

individual researchers, as they feel the need to personally make up for the lacking resources and engage themselves in auditing practices.

Some reformers have attempted to address the infrastructural gap by building their own version of a dedicated workforce. In her current role as editor of the prominent psychology journal *Psychological Science*, Simine Vazire has created a new category of editors – statistics, transparency, and rigour (STAR) editors – whose work takes place alongside and after that of traditional academic editors who manage peer review and make decisions on manuscripts.[69] The STAR editors do the laborious work of checking the manuscript against its preregistration, checking that the analysis is computationally reproducible, and so forth. This is, however, an institutionally fragile model because of the extent to which it relies on volunteer labour. Reflecting on these and other limitations of preregistration, one recent opinion piece from political scientist Rose McDermott concluded that 'preregistration represents a form of virtue signalling that is more performative than actual' since overburdened editors and reviewers are unlikely to monitor compliance with a preregistered plan in most cases.[70]

Finally, McDermott's characterization of reformers as engaging in 'virtue signalling' and 'crush[ing] creativity' demonstrates how common critiques of regulation take on a different character when the regulatory ethos moves into other areas of science.[71] Regulatory agencies have long been subject to critiques that they are slow, bureaucratic and overly cautious in their approaches. In the United States, for example, the FDA faced increasing pressure in the 1990s and 2000s from critics who asserted that the agency's requirements were 'stifling innovation'.[72] The Critical Path Initiative, introduced in 2004, aimed to 'modernize' the FDA's processes by incorporating omic and imaging biomarkers into clinical trial designs.[73] Somewhat ironically, while critiques of the scientific status quo have pushed academic science more towards the regulatory ethos, critiques of FDA bureaucracy have pushed this key regulatory agency away from it. Bayesian adaptive clinical trials for cancer therapeutics in the 2010s and accelerated approval of Alzheimer's drugs in the 2020s are both examples of how the agency attempted to incorporate flexibility into its processes that would allow for the direction of a clinical trial or the evidence base for a new drug to evolve over time rather than being strictly specified in advance.[74]

More important for our argument, however, is how these critiques make visible existing scientific identities, values and practices that conflict with the regulatory ethos – that is, they make visible other forms of scientific life. When NIH deputy director Larry Tabak first presented the NIH's plans for reproducibility reforms to the Advisory Committee to the Director in December 2013, he paused to joke about the 'number of frowns he saw around the table', and that he was well aware that many people saw some of the proposed interventions as 'constraining innovation'.[75] Members of the Advisory Committee reflected back his concerns. One academic biologist made a 'plea' to the NIH to ensure that there is good

[69] Association for Psychological Science, 'Vazire outlines goals for transparency, diversity in psychological science', at www.psychologicalscience.org/publications/observer/simine-vazire-psychological-science.html (accessed 10 January 2025).

[70] Rose McDermott, 'Breaking free: how preregistration hurts scholars and science', *Politics and the Life Sciences: The Journal of the Association for Politics and the Life Sciences* (2023) 41(1), pp. 55–9, 55.

[71] McDermott, op. cit. (70), p. 57.

[72] Janet Woodcock and Raymond Woosley, 'The FDA Critical Path Initiative and its influence on new drug development', *Annual Review of Medicine* (2008) 59(1), pp. 1–12.

[73] Woodcock and Woosley, op. cit. (72).

[74] Donald A. Berry, 'Adaptive clinical trials in oncology', *Nature Reviews Clinical Oncology* (2011) 9(4), pp. 199–207; Billy Dunn, Peter Stein and Patrizia Cavazzoni, 'Approval of Aducanumab for Alzheimer disease: the FDA's perspective', *JAMA Internal Medicine* (2021) 181(10), pp. 1276–8.

[75] Advisory Committee to the Director, 'Advisory Committee to the Director (ACD) December 2013 – Day 1', NIH VideoCast, at https://videocast.nih.gov/watch/a741b831-d5db-11f0-9cf9-12c45c580ad9 (accessed 2 March 2026).

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

evidence for the efficacy of new requirements before imposing them, because the scientific community is already 'drowning in endless bureaucracy'. Others voiced overlapping concerns that they 'did not want to take away what is unique about academia' and that we need to recognize the important role of intuition and not just 'throw the old away.' They pointed to important historical discoveries in biology built on close observation rather than pre-specified plans to underscore the importance of other styles of work.

The only member of the committee to speak positively about the proposed interventions was a senior executive at GlaxoSmithKline. He saw the differences in reactions – and the differences in academic and industry research – as a 'cultural divide'. Over the fifteen- to twenty-year life cycle of developing a drug, he explained, many scientists will be involved in the process, and so practices aimed at providing continuity within a company were important. These practices might look bureaucratic or inflexible to outsiders, but, he argued, they were a fundamental part of the culture of industry. Implicit in his comments was the suggestion that this industry culture was also the more desirable culture. He spoke with frustration about how academics faced 'no consequences' when they failed to provide enough information for others to replicate their work, and if that work eventually proved to be irreproducible academics could 'even write a paper about why the previous paper was wrong'.

As we have shown in this section, the practices being promoted by reproducibility reformers bear strong similarity to the practices used by regulators to shape knowledge produced for regulatory purposes. But this exchange in the NIH Director's Advisory Committee meeting demonstrates that it is not simply a set of practices that is being promoted; it is an ethos that encompasses both practices and a set of beliefs about why those practices represent a right, ethical and proper way to be a scientist. In a proposed list of paperwork requirements, the pharmaceutical industry executive saw a familiar 'culture', while the academic scientists saw not just the pragmatic problems of increasing bureaucracy but a deeper mismatch between the spirit of what they saw represented in the reforms and what they saw as unique about their own ways of being a scientist.

Conclusion

Our characterization of the regulatory ethos offers a new analytic framework to examine the intertwinement of moral, technical, institutional and epistemic elements in contemporary evaluations of science. It allows us to draw together several points of critique that both academic scientists and analysts of science have raised regarding specific features of the regulatory ethos (that is to say, the preference towards planability, uniformity, accountability and reliability) and their portrayal as generally better, more trustworthy ways of doing science.[76]

Academic scientists' objections to reproducibility reforms are often loosely articulated, relying on vague assertions about the importance of creativity or the uniqueness of academic culture. Scholarship from historians, philosophers and STS scholars can help provide a better vocabulary for identifying what might be lost in the turn towards a regulatory ethos and what the alternatives are. Several scholars have already begun pursuing this project, pointing out that there are many ways of generating knowledge and each has its strengths and weaknesses. Sarahanne Field and Maartin Derksen, for example, argue that rather than leaning into a narrow form of objectivity to ground psychology, the field might be better

[76] Sabina Leonelli, 'Rethinking reproducibility as a criterion for research quality', in Luca Fiorito, Scott Scheall and Carlos Eduardo Suprinyak (eds.), *Research in the History of Economic Thought and Methodology*, vol. 36, Leeds: Emerald Publishing Limited, 2018, pp. 129–46; Bart Penders, J. Britt Holbrook and Sarah de Rijcke, 'Rinse and repeat: understanding the value of replication across different ways of knowing', *Publications* (2019) 7(3), p. 52.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

served by using reflexive practices that draw more attention to subjectivity as a way of interrogating bias in psychological research.[77] Penders argues that the dichotomous distinctions created by reformers could be complicated by other, cross-cutting distinctions that would allow for 'the maintenance of epistemic and methodological plurality'.[78] David Peterson and Aaron Panofsky argue that the claim that metascience reforms improve efficiency could be challenged since the goal that reformers want to more efficiently progress towards is often ill-defined.[79]

Our definition of the regulatory ethos is likewise aimed at demonstrating that the ethos represented by these reforms is not synonymous with the ethos of science writ large, and that the reference to an ethos comes in general with a virtue-theoretical focus that might not be sufficient to evaluate whether the epistemic and pragmatic goals of such virtuous science are achieved. By positioning each feature against a contrasting feature that is also recognizably scientific, we aim to show that the conflict is one not merely between subjectivity and objectivity, but between different forms of objectivity. And by describing this ethos as regulatory, we aim to connect long-standing discussions about the strengths and weaknesses of regulatory science to this more contemporary conversation about scientific reform, thereby opening up more avenues for scholarly inquiry. Identifying similarities between the regulatory ethos and reproducibility reforms, for example, raises questions about how practices born in regulated sectors will function when they are stripped of their law-enforcing and institutional power.

In our view, the most important insight arising from a sustained examination of the regulatory ethos is the danger of the slippage from validation to validity. Within the framework of the regulatory ethos, the focus is on securing validation processes that are attuned to reproducibility and reliability rather than treating validity as a tool of critical inquiry into the premises of testing and the transferability of its results. The history of regulatory science provides many instructive examples of how even well-executed, 'gold standard' approaches can fail to produce findings that correspond with meaningful, real-world phenomena. For example, Steven Epstein's case study of HIV/AIDS clinical trials showed that the FDA's insistence on 'clean' clinical trials ironically resulted in substandard data, in multiple senses.[80] Patients forced into a placebo-controlled trial were rumoured to have smuggled drugs from other countries or pooled their pills with fellow study patients to maximize their chances of receiving some active drug, thereby compromising the study's results. Additionally, the trialists' insistence on recruiting 'treatment naïve' patients meant that the data produced had less value for predicting how the drugs might function in patients with multiple illnesses and/or taking multiple medications – as one group advocating for the use of 'real-world evidence' put it, 'the internal validity attained in these trials is often achieved at the expense of uncertainty about generalizability'.[81] While proponents of the regulatory ethos may see this more innocuously as simply an ordering of priorities – a desire to get the 'procedural hygiene' of their field in shape before moving on to deeper, less tractable problems – the HIV/AIDS case study and the cases presented in this

[77] Sarahanne M. Field and Maarten Derksen, 'Experimenter as automaton; experimenter as human: exploring the position of the researcher in scientific research', *European Journal for Philosophy of Science* (2020) 11(1), article 11.

[78] Penders, op. cit. (9).

[79] David Peterson and Aaron Panofsky, 'Arguments against efficiency in science', *Social Science Information* (2021) 60(3), pp. 350–5.

[80] Steven Epstein, *Impure Science: AIDS, Activism, and the Politics of Knowledge*, Berkeley: University of California Press, 1996.

[81] Rachel E. Sherman, Steven A. Anderson, Gerald J. Dal Pan, Gerry W. Gray, Thomas Gross, Nina L. Hunter, Lisa LaVange, Danica Marinac-Dabic, Peter W. Marks, Melissa A. Robb, Jeffrey Shuren, Robert Temple, Janet Woodcock, Lilly Q. Yue and Robert M. Califf, 'Real-world evidence: what is it and what can it tell us?', *New England Journal of Medicine* (2016) 375(23), pp. 2293–7, 2294.

https://doi.org/10.1017/bjt.2026.10041 Published online by Cambridge University Press

issue of *Themes* illustrate that, in some cases, adopting this ethos can actively work against validity.[82]

Acknowledgements. We would like to thank Angela Creager and the members of the Validation and Regulation Working Group, audiences at the Institute for Studies of Science and the Philosophy of the Life Sciences Colloquium at Bielefeld University, the Max Planck Institute for the History of Science, the History of Science Seminar at Leopoldina, the Medical Museion in Copenhagen, the History of Medicine Seminar at Cedars-Sinai, the History of Science, Technology and Medicine programme at the University of Wisconsin–Madison and two anonymous referees for very helpful feedback. Parts of our research have been funded generously by the Max Planck Society, Bielefeld University, the Harvard Radcliffe Institute for Advanced Study, and the University of Wisconsin–Madison.

Competing interests. We have no competing interests to declare.

[82] Bart Penders, 'Renovating the theatre of persuasion: ManyLabs as collaborative prototypes for the production of credible knowledge', *MetaArXiv* (2024), at https://doi.org/10.31222/osf.io/vhmk2 (accessed 1 March 2026).

Cite this article: Nicole C. Nelson and Lara Keuck, 'The regulatory ethos in science,' *BJHS Themes* (2026), pp. 1–19. https://doi.org/10.1017/bjt.2026.10041