

The book focuses on how research evaluations transform academic disciplines. Focusing on four social sciences (anthropology, economics, politics and international relations, and sociology) it presents evidence of the effects of research evaluations on the careers and knowledges produced by UK-based academics. Combining computational methods, quantitative longitudinal models of labor mobility, and conversations with scholars from across Britain into an approach inspired by Burrawoy's extended case method, the book shows how the logics of this family of research assessments slowly shift the intellectual work and careers of social scientists over time. Anchored on disciplinary rationalities of what constitutes "relevant science" in each field, these exercises change academic disciplines toward greater levels of organizational and epistemic homogeneity. With time, evaluation after evaluation, exercise after exercise, the knowledge we produce and the spaces we inhabit became more similar, more isomorphic.

In searching for the roots of this homogeneity, *The Quantified Scholar* points at the role that our shared vocations play in amplifying practices and logics of evaluation. The chains we place on knowledge may be forged by others, but they are partly inspired by what we collectively accept as scholarly practices. In thinking about how to make knowledge that is more adaptive to our future challenges, *The Quantified Scholar* advocates for rethinking evaluation from the bottom up, shifting from a way of thinking about academia that encodes problematic structures from the past to a collective way of organizing work and knowledge oriented at a politics of everyday solidarity.

DOI: 10.1111/1468-4446.13048_2

The trials and tribulations of research evaluation: Quantification to the rescue?

John Holmwood 

Centre for Science Technology and Society Studies, Institute for Philosophy, Czech Academy of Science, Prague
Email: John.Holmwood@nottingham.ac.uk

The UK Research Excellence Framework (REF; formerly the RAE, or Research Assessment Exercise) takes place at 5–7 year intervals. It is the means through which block-grant public funding of university research is distributed (so-called QR funding). This exists as a major source of university research funding alongside competitive funding for specific research projects distributed by research councils and other bodies. In its various iterations the REF has been a feature of UK academic life for over 30 years. Its criteria have been incorporated into annual performance reviews of departments and academic staff, such that it can properly be described as part of the everyday experience of all "research active" (as the jargon has it) academics on teaching and research contracts in the UK.

Over the years the REF has elicited considerable commentary from specialist science policy researchers concerning the validity of its measures and "perverse effects" (e.g., Pinar & Horne, 2022) and from individual academics commenting on its consequences for academic well-being (e.g., Gill, 2016). Some of these issues were addressed in the Stern Review of 2016, which required all research active staff to be submitted with at least 1 publication and an average of 2.5 publications for each person submitted overall (compared with 4 publications per submitted individual in earlier REF exercises, with universities deciding which staff to put forward for evaluation). This was intended to prevent "gaming" and to reduce the impact of the REF on academics, especially early-career academics. Notwithstanding, the juggernaut rolled on.

While this piece was being written, a new framework was announced for REF 2028 (Research England, 2023). Among its proposals is that "outputs" will be separated from individuals. A staff "roll" for each University's subject Units of Assessment (UoAs) will be calculated from a register created centrally by the Higher Education Statistics Agency (HESA). This will be used to calculate a quota of full time staff (FTEs) for an UofA where a submission would

be required to submit outputs equivalent to $2.5\times$ the assigned FTEs. Individuals associated with the UofA need not have an output submitted against their name and there would be no upper limit in the number that could be submitted by any individual. Not all members of staff in the quota will need to be in academic research roles.

As with Stern, the immediate academic response (on twitter) has been to welcome the development which, it is held, would reduce the pressure of REF on individual academics. Of course, the separation of “academic” and “output” is only true at the aggregate level of the REF itself—the association of individuals with their outputs necessarily remains, not least for the management of the REF at individual institutions. The importance of Juan Pablo Pardo-Guerra’s book is that it is specifically an examination of the local effects of the REF process.

The UK is not alone in having a formal system of accountability for publicly-funded research in universities. In its all-encompassing nature, the REF is an outlier, perhaps, but it is one that shows most clearly the way in which a rigorously applied audit system intended to create a competitive environment for university research has resulted in a vast bureaucratic system that empowers university managers. In this important, powerfully argued and well-researched book, Pardo-Guerra provides a novel analysis that cuts through the plethora of other commentaries. He addresses the consequences (intended and otherwise) of formal research evaluation exercises for disciplinary formation in the social sciences. To my knowledge, it is the first book to be written on the impact of research evaluation on the structure and practices of academic disciplines. It demands careful attention and reflection because its lessons have a wide significance and will help focus the mind on what is proposed for REF 2028.

In these respects, its approach is less about an evaluation of research assessment as a policy (although it necessarily has important messages here) and more about the application of a critical and well-evidenced sociology of science. The book is focused on four REF sub-panels out of the 11 gathered under Main Panel C (clusters of sub-panels were first introduced in RAE 2001). These four are anthropology and development studies, economics and econometrics, politics and international studies, and sociology. It constructs a series of data-sets of outputs submitted and bibliographic information derived from those outputs in order to analyze changes in the intellectual profile of submissions across the 2001, 2008 and 2014 REF exercises. While the book draws much of its data from the 2014 exercise (in which the author was part of the sociology submission at LSE), it also includes interviews conducted in its wake. It was published after the results of the Covid-delayed REF2021 exercise were announced in March 2022.

Responses to the latter have been muted compared with those of previous exercises and have involved two main responses, both themselves artifacts of the REF. The first is a push to greater efficiency by the administering body at UKRI, calling for a streamlining of the process by shifting the evaluation of the research environment component of the REF to university-level, rather than that of individual subject submissions. The second is to use the increase in the proportion of scores at the highest level of international excellence (something that happens in each REF-cycle!) to call for increased government spending on university research on the basis of its demonstrable value for money. This is considered especially urgent in the light of tardiness by the government in putting in place post-Brexit funding arrangements to cover the loss of revenue from withdrawal from EU research funding schemes.

This has affected the reception of Pardo-Guerra’s book in the UK, where it risks spoiling this particular party. Interestingly, the book provides an explanation of this reception. The REF is a co-produced and self-legitimizing system of evaluation. It is reproduced through the active participation of those subject to its devices. It is both highly centralized (nationally through Research England on behalf of UKRI) and within each university and its research governance system and dispersed through subject panels, departments/or UofAs and in the practices of individual academics. Each level involves multiple forms of buy-in—including by professional associations which make recommendations for membership on the subject panels. The general consequence of the REF, as Pardo-Guerra shows, has been to produce a series of what Andrew Abbott (2001) calls “self-similar structures” of research practice and management.

But before I address these aspects, let me say something about the distinctiveness of Pardo-Guerra’s methods and the data sets he constructs through which these consequences of the REF are displayed. These necessarily have their limitations, but they are similar to those of any path-breaking study. Certainly, there are gaps to be filled in, but it is hard to gainsay the findings. Or, at least, gainsaying involves the claim that the findings might have been different if, say, books were more securely identified in citation data. Or, that it is no longer relevant given the changes introduced between REF

2014 and REF 2021. Or, that the judgments revealed in the REF correspond with an underlying reality, one, moreover, that is confirmed by the judgments of colleagues, where the latter correspond with an “inner truth” with more validity than any crude metric. The “REF-reproduction” machine operates in just the way in which Pardo-Guerra describes.

Notwithstanding any disciplinary self-understanding of scholarship and research as a form of critical engagement, one consequence of the REF as co-produced has been the co-production of resistance to criticism of how it has shaped the research of those most actively caught up in it. Mostly, criticism of it has been individualized in terms of academic wellbeing or, if expressed collectively, it is directed at its contribution to increased managerialism within the university (alongside other audit mechanisms); yet, that managerialism is itself also co-produced.

These responses, we should understand from Pardo-Guerra, are a necessary feature of the REF itself. After all, as I have indicated, the REF determines the distribution of QR funding which is crucial to the fate of departments. In addition, position in a rank-order of outcomes has become increasingly important as a means of recruiting students in the new marketplace of student applicants bearing government organized income-contingent loans that was put in place in England after the ending of direct funding of undergraduate places in 2011. The REF panels have to be populated in order to produce these “positional goods” that have become the inescapable features of academic life and we should be grateful to those of our colleagues who have taken on the mantle. All these aspects secure compliance, or, at least, grudging acceptance (Bottero, 2022).

The REF is self-referential as well as self-reproducing. Each REF exercise shows an increase in the international significance of its submitted outputs. At different times, reforms have been suggested to modify its character—the Stern Review has now been superseded by FRAP 2023 (the Future Research Assessment Programme)—but they only seem to have reinforced its existing characteristics. This becomes clear when we consider the consequence of REF 2021, which accentuated tendencies evident in Pardo-Guerra’s analysis of developments up to REF 2014. For example, in REF 2021, despite there being a 46% increase in FTE researchers submitted following the Stern reforms (Sociology had a 56.8% increase), the proportion of outputs receiving the highest scores also increased. Seeming differences in judgments across panels are attributed to real differences in quality, while the rectification over time of apparent anomalies is similarly “real,” attributed to improvement across the REF-cycle, so the increase in 4* outputs in Sociology, say, from a relative low-level of 21.1% in 2014 to 31.8% in 2021 can be explained by an improvement in the research environment in the submissions (all data are taken from the Main Report of Panel C). REF 2028 will potentially have a similar number of staff FTEs assigned to the REF, but only a sub-set of them will have their outputs submitted. Since local decisions about the submission will be based on assessments of likely score, we can anticipate a further increase in 4* outputs.

While Pardo-Guerra’s book is focused on data from REF 2014, his qualitative interviews were conducted in relation to preparations for REF 2021. In that sense, they show that the Stern reforms—the requirement that all staff with research as part of their contract should be included in a submission, with the requirement that each individual submitted should have at least 1, rather than 4, items in a ratio of 2.5 for each FTE submitted—did not loosen the turn of the REF audit screw. Moreover, these interviews show that there was significant buy-in into the logic of the REF, down to departmental level in the form of the locally self-similar structure—that is, on the part of research directors and REF leads, and in terms of the incorporation of REF criteria into promotion criteria. This will not change under the new arrangements for REF 2028. Whether or not an individual member of an UofA has outputs selected for submission will be locally consequential.

Of course, this process is not straightforward for all disciplines. Of the four subject areas examined by Pardo-Guerra the responses by economists were most aligned with the assumptions of the REF, and sociologists and anthropologists the least (poignant, then, that Economics lost most from the application of the funding algorithm to the results; Royal Economic Society, 2023). One of the most striking findings, however, is that of a convergence of all disciplines on a pattern best exemplified by Economics at an earlier point in the development of the REF, where UofA/departments have become more similar to each other in terms of the representation of the discipline through a set of common fields and topics. Pardo-Guerra shows this to be a product of the REF as each cycle is planned for by universities through their research strategies and REF-readiness exercises. It is also reproduced by individuals in terms of their career movements and the hiring strategies of departments.

Paradoxically, as the “international standing” of the disciplines increases (as measured by the increasing proportion of REF scores at 4* over the different REF cycles), so the capacity for innovation is reduced—that is, if we assume that to be derived from self-directed research within an intellectually diverse field. Research has become a matter of central planning within universities. Is there a UK university without a centrally-defined research strategy? These are set out to identify research excellence and “innovative” work on “global challenges.” Nearly all universities promote the same challenges: environment, health, sustainability, security, justice, which are also reproduced in the research funding calls of UK research councils.

The proposals for REF 2028 will reinforce these tendencies by creating a hierarchical structure where an institutional-level statement of “people, culture and environment” has to incorporate subject-level statements, and submissions of outputs have to incorporate a statement of alignment with the institutional-level evidence statement. In effect, this tightens the grip of institutional-level research management over the direction of research within UofAs.

This convergence is also reinforced by the “impact agenda” which directs research toward its co-production with “users” and “beneficiaries” (Holmwood, 2011). It is applied both to research supported by research council funding (the “outputs” of which will later appear in REF submissions) and to QR funding via the REF. In the latter case, the requirement is to submit impact case studies for evaluation, something first introduced in REF 2014. Each UofA was required to submit 2 case studies for a submission of up to 15 FTE research active staff, with an additional case study required for each further 10 staff submitted; to be modified by a less stringent requirement for REF 2028. Pardo-Guerra was not able to address the role of impact in his study, but it would seem to reinforce the tendencies that he is otherwise describing, and it potentially represents a broader shaping of disciplines than the one he is describing within the four sub-panels he addresses.

In an earlier discussion of science policy, Nowotny et al. (2001) coined a distinction between two modes of knowledge production, Mode 1 and Mode 2 knowledge. The latter consists in interdisciplinary, applied problem-based knowledge, which they contrast with discipline-based knowledge, Mode 1. Where Mode 2 involves the external beneficiaries of research drawn into the research process as its co-producers, Mode 1 is described in terms of disciplinary hegemony and internal audiences of peers. This distinction is carried into the organization of REF sub-panels, which each include academic peers alongside user members.

More significantly, however, the sub-panels are themselves divided in ways that map onto the distinction between Mode 1 and Mode 2 knowledge. For example, Economics and Econometrics would seem to be primarily organized in terms of Mode 1 knowledge, while the Business and Management Studies sub-panel includes the former’s Mode 2 complement. A similar situation applies to Sociology and its complementary Social Work and Social Policy sub-panel. Indeed, there are more sociologists submitted under the latter sub-panel than are submitted to the Sociology sub-panel, which suggests that Pardo-Guerra’s description of convergence among sociology submissions might be further reinforced were sociology departments submitting to Social Work and Social Policy also included in the submission. Anthropology, in contrast, is included in a sub-panel alongside its complementary applied area, Development Studies. In addition, of course, there are a number of other sub-panels that represent Mode 2 knowledge, including Education and Sport and Exercise Sciences, Leisure and Tourism. Just over half of the 21,450 FTEs submitted under Main Panel C were in the sub-panels for Business and Management Studies, Education, Social Work and Social Policy and Sport and Exercise, Leisure and Tourism, compared with just over one fifth for the four panels studied by Pardo-Guerra, showing the predominance of Mode 2 knowledge in UK social science.

It is interesting to note that an earlier path-breaking study by Whitley (2000 [1984]; see, Holmwood, 2014 for discussion) on the structure of disciplines includes factors associated with the rise of Mode 2 knowledges—for example, he constructed a typology of disciplinary formation organized in terms of high or low task uncertainty (relating to the agreement or otherwise on methods to be used) and strategic uncertainty (relating to agreement or otherwise on problems deemed important for study), and high or low functional and strategic dependence (relating to the tight, or loose, “coupling” among researchers and their problems). The closer a discipline was to high uncertainty/low dependence the more it was likely to be open to external factors in the determination of its fields.

The REF, itself, represents the most important of external, determining factors in the organization of UK social sciences (alongside student recruitment with which it is connected via the role of the REF in the construction of prestige rankings). We might expect that this would change Economics as a discipline, too, such that the convergence of other disciplines with it under the pressure of the REF has also included movement within Economics. Perhaps this is masked by the fact that it is not so much that Economics, as defined in REF terms, has become less like itself, as that it has simply become *less* in the simple sense that there are now more economists within the Business and Management Studies submission, just as there are more sociologists in the Social Work and Social Policy submission than in the Sociology submission. The changed methodology between REF 2014 and REF 2021, where all staff have to be submitted (albeit with fewer outputs) make this difficult to substantiate across Panels, but the Main Report for Panel C shows that the number of staff in the Economics and Econometrics sub-panel increased by 21.7% (the smallest rate of increase) compared with 99.8% for Business and Management Studies (which had the largest rate of increase).

These tendencies are further reinforced by the fact that the REF requirements for submissions are the same for all sub-panels. Applied subject areas are aligned with the Mode 2 understanding of knowledge production that underpins the general idea of impact case studies, but the Mode 1 disciplinary subject areas are enjoined to demonstrate their contribution to external users. The “fact” that they have “succeeded” in matching them becomes part of the reproduced boosterism of the REF process itself. In this way, the REF has changed the structure of the social sciences and, at the same time brought about a convergence of subject areas on the characteristics of Mode 2 knowledge.

This has had the paradoxical consequence of laments about whether or not “interdisciplinary” knowledge is properly recognized within the REF, when most subject panels have become interdisciplinary and applied in character. The REF generates a lot of discussion of the problem of “interdisciplinarity” and how it is evaluated, where this is treated as interdisciplinary across sub-panels. For example, Main Panel C includes Business Studies as well as Economics, and Social Work and Social Policy as well as Sociology. Is Business Studies a discipline? Is Social Policy? Or Sport, Leisure and Tourism? Indeed, it would be more accurate to argue that it is *disciplinary* knowledge that is not properly recognized and the real problem is the absence of work on the boundaries of *disciplinary* concepts, such as those of sociology and economics, say, or sociology and politics (Holmwood, 2010). Certainly, the cross-referral of outputs from one sub-panel to another is more about the way that universities have aggregated academics into a submission than it is about the assessment of intrinsically cutting-edge work at the boundaries of disciplines.

In the “REF-speak” of future REF 2028 the self-referential nature of the process becomes both the definition and guarantee of excellence: “the funding bodies agreed that supporting a healthy research culture should be an underpinning principle of the REF, which influences all aspects of its design and conduct. At every stage of policy development, the funding bodies have considered the impact of their decisions on research culture within UK universities” (Research England, 2023, p. 20). We now have a convergent research culture determined by senior managers and described as meeting the interests of a community of researchers that is claimed to be intellectually diverse but whose members are increasingly self-similar.

As Pardo-Guerra shows, changes in the structure of disciplines and the increased similarity of departments also have consequences for equality and diversity in staffing. This is so despite the fact that the REF protocols require Universities and UofA to show that they have effective policies addressing such issues. Thus, disciplinary consolidation of sociology favors male academics over female and white British over ethnic minority academics. If all departments are moving to a similar spread of topic fields might this be a factor in the failure to overcome inequalities in recruitment and promotion? For example, are women and ethnic minorities more concentrated in terms of topic fields? How might we compare economics and sociology in this regard, where the former discipline is both more “internationalised” and, at the same time, more male-dominated? Significantly, the proposals for REF 2028 mean that, in the view of its architects, with the link between individuals and outputs severed, it will no longer be necessary for there to be statements about the consequence of individual staff circumstances. Of course, staff will continue to experience different circumstances, and experience them differentially by gender and ethnicity.

Finally, Pardo-Guerra addresses how the formal processes of the REF are accentuated by “mock” exercises within institutions (which are, of course, also related to recruitment practices and annual reviews). It looks as if this has been exacerbated, not ameliorated, by the changed methodology for REF 2021, given that his interviews were about preparation for the latter. It will not be different in REF 2028. Indeed, the implications of his analysis are that the problem does not straightforwardly lie with the REF as such, but in the ways in which academics have reinforced its worst aspects by their own competitive behavior, whether for mobility and promotion or for the reflective glory of a highly-ranked department. This is why the new criteria for REF 2018 are unlikely to change anything.

Pardo-Guerra shows how the quantification of scholarly activity through the REF has produced what might be termed an iron cage of research organization, but the implication is that the bars have been forged within academic practices. He poses the question can this be mitigated by the self-organization of scholars determined to act in a collegial manner? Indeed, he directs us toward this as an objective—including the possibility of a new transformative and transgressive practice. He invokes what we might have learned from our experiences during Covid. Have we missed the post-Covid moment? Certainly, REF 2021, although disrupted by Covid, has elicited little active response directed at changing it. It might be its massive cost in time and effort, rather than a re-found collegiality, that is its downfall.

The solution might seem counter-intuitive, but the implication of Pardo-Guerra's analysis is that it is not so much quantification that is the problem, but the self-reproducing, co-produced nature of the REF. A radical reform of the REF was on offer in the Stern Review, when the possible alternative was the use of metrics. This was widely attacked on the (legitimate) grounds that metrics would be unfair when applied to individual outputs and for finely-graded judgments. In hindsight, what was intended to maintain collegiality—peer judgment rather than metrics—has undermined it. However, as a way of reaching aggregated judgments metrics are much less problematic and can be modified to the needs and characteristics of different subject areas. In fact, Pardo-Guerra's own book shows the merits of quantification. Indeed, metric data is automatically generated by our own publication and citation practices and is already available to research managers via automated dashboards. The horse has bolted despite our attempts to lock the stable door.

Moreover, there is no need for fine judgments. The distribution of QR money could be done in relation to broad thresholds (as was done in the earliest iterations of the RAE). This would also undercut the construction of rank orders. We already grade our undergraduates in this way—first, upper second, lower second, etc.—where the vast majority are in the top two bands. That is a lot like the underlying reality of the REF! The outcome would be a potentially fairer distribution of QR income with less effort spent on mimicking and modeling the REF and more time available for research and scholarship. Might quantification save the quantified scholar? *We have nothing to lose but our chains* (and an academically-defined research culture to gain).

ORCID

John Holmwood  <https://orcid.org/0000-0002-3440-3524>

REFERENCES

- Abbott, A. (2001). *Chaos of disciplines*. University of Chicago Press.
- Bottero, W. (2022). Grudging acts. *Sociology*, 57(3), 533–551. <https://doi.org/10.1177/00380385221104017>
- Gill, R. (2016). Breaking the silence: The hidden injuries of neo-liberal academia. *Feministische Studien*, 34(1), 39–55. <https://doi.org/10.1515/fs-2016-0105>
- Holmwood, J. (2010). Sociology's misfortune: Disciplines, interdisciplinarity and the impact of audit culture. *The British Journal of Sociology*, 61(4), 639–658. <https://doi.org/10.1111/j.1468-4446.2010.01332.x>
- Holmwood, J. (2011). Viewpoint – The impact of ‘impact’ on UK social science. *Methodological Innovations Online*, 6(1), 13–17. <https://doi.org/10.4256/mio.2010.0025>
- Holmwood, J. (2014). Sociology: The impact of external structure, policy and financing. In J. Holmwood & J. Scott (Eds.), *A handbook of British sociology*. Palgrave-Macmillan.
- Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-thinking science*. Polity.

- Overview Report by Main Panel C and Sub-Panels 13 to 24. May 2022. <https://www.ref.ac.uk/media/1857/mp-c-overview-report-final.pdf>
- Pinar, M., & Horne, T. J. (2022). Assessing research excellence: Evaluating the research excellence framework. *Research Evaluation*, 31(2), 173–187. <https://doi.org/10.1093/reseval/rvab042>
- Research England. (2023). Research Excellence Framework 2028: Initial decisions and issues for further consultation (REF 2028/23/01). <https://www.jisc.ac.uk/future-research-assessment-programme/initial-decisions>
- Royal Economic Society. (2023, January 12). RES Statement on funding allocations arising from REF2021. <https://res.org.uk/res-statement-on-funding-allocations-arising-from-ref2021/>
- Stern Review. (2016). Research Excellence Framework (REF) review: Building on success and learning from experience. <https://www.gov.uk/government/publications/research-excellence-framework-review>
- Whitley, R. (2000 [1984]) *The intellectual and social organization of the sciences*, 2nd ed. Oxford University Press.

DOI: 10.1111/1468-4446.13048_3

The role of the social sciences in the quantified university. Response to *The Quantified Scholar*, by Juan Pablo Pardo-Guerra

Sarah de Rijcke 

Centre for Science and Technology Studies (CWTS), Leiden University, Leiden, The Netherlands
Email: s.de.rijcke@cwts.leidenuniv.nl

Calls for more responsible forms of research assessment have become a prominent feature of contemporary discourse on addressing perverse effects of performance measurement and systemic forms of exclusion and inequity in the academy. Often, however, such calls are grounded in narrow understandings of how epistemic practices, accountability measures, and society relate. Equally often, such calls do not address the root causes of the emergence of singular mechanisms for knowing quality in academic settings.

As part of larger reform initiatives in contemporary academia, many scientific disciplines have recently been structured along market economic principles in order to simultaneously encourage competitiveness and responsiveness to societal needs (Nowotny et al., 2001). These shifts in the governance of science include significant changes in the funding structures (Whitley, 2010), increasing formalization of scientific work by way of project management (Fowler, 2015) and the integration of multiple assessment mechanisms at various levels of the system (de Rijcke, 2017; Power, 1997; Whitley & Gläser, 2007).

In his new book, Juan Pablo Pardo-Guerra delves into the impacts of all this evaluating and measuring on the social sciences in the United Kingdom. The book demonstrates that a dominant thinking in terms of market economics and increasingly standardized and metricized research evaluations have led to reduced diversity and increased conformity within UK social science disciplines. These evaluations have rewarded social scientific work that aligns closely with established disciplinary norms, thereby reinforcing publication hierarchies and longstanding forms of academic prestige. The book also calls for the cultivation of new forms of vocational solidarity, to challenge and disrupt entrenched inequalities within academia.

One immensely commendable aspect of the book is that it incorporates an extensive range of literature, spanning from sociology of science to science and technology studies, to anthropology and history and philosophy of science. Additionally, the author also skilfully employs various methodologies: longitudinal data, statistical and computational models, interviews, and oral histories are seamlessly integrated. It makes the book a truly interdisciplinary and mixed method endeavor, which is precisely what this challenging subject matter needs. Not only does it enhance our understanding of the interplay between evaluation, quantification, and scholarly work, but it also sheds light on the epistemic implications of quantification. Over and above, Pardo-Guerra incorporates his own lived experiences and