1. Introduction

Rob van Gestel and Andreas Lienhard

1. QUALITY OF ACADEMIC LEGAL RESEARCH IN EUROPE

Who are the best legal scholars in Europe in different fields of law? Which journals are considered the best in Europe and what sorts of assessment methods do editorial boards and publishers apply when evaluating manuscripts? These are but a few of the questions that are difficult to answer for legal academics, university managers and funding bodies. To outsiders, these kinds of questions might seem trivial, because in most other (social) sciences many scholars know each other's h-index, there are official rankings of journals and publishers, and editorial boards are quite clear about the standards they apply for (single or double blind) peer review. In law, however, all this is different. Not only do legal academics in Europe publish a wide variety of articles, essays, books, commentaries, case notes and so on, in a broad range of languages, about a wide variety of national legal systems; but unlike scholars in the hard sciences, they also address multiple audiences. The readership of legal scholars ranges from other academics to courts, solicitors, legislators and so on. Without realizing it, the absence of uniform evaluation practices for academic legal publications may have unexpected consequences.

To give one practical example. In an editorial for the *European Journal of International Law* (*EJIL*), Joseph Weiler discusses the current evaluation practice in the field of law. He signals that for a number of years, the *EJIL* has been one of the highest-ranked European law journals in the US Washington and Lee (W&L) journal rankings. On the one hand, that looks like something to be proud of. On the other hand, the *EJIL*, as a double-blind, peer-reviewed journal, does not even come close to the top of the list of journals in the W&L rankings and is surpassed even by some student-edited journals. However, taking a deeper look at the ranking methodology reveals that there is a perfectly logical explanation for this: the W&L rankings mainly count citations in US law reviews, whereas most *EJIL* citations are probably found in other European journals. The trouble is that there is no central European citation

database for legal publications, which means that even English language law journals such as the *EJIL* will probably continue to lag behind US journals, for which there is such a database.[1]

This example shows why, in an increasingly globalized legal world in which competition for funding is becoming increasingly intense and rankings have gained considerable influence, not accurately showing the impact of one's research to the outside world might have serious negative consequences. At the same time, though, the example reveals how unreliable existing assessment methods such as rankings can be. Simply because of how citations are counted, European law journals will never come out on top of the W&L rankings − even if we leave aside for a minute the fact that most European law journals are not published in English, which hinders accessibility for foreign scholars and thereby limits their impact. This could, of course, be seen as a warning sign to establish European law journal rankings and a European citation database. Although this sounds sensible, it is far from clear whether this would be such a good idea. After all, how does one compare the quality of law journals written in different languages for various audiences (professionals and academics) by authors from different backgrounds (scholars and practitioners) and different legal cultures (e.g., common law versus civil law)?[2] Moreover, to what extent do rankings of law journals actually measure the quality of research; and is putting more emphasis on competition between journals and academics necessarily the best approach when the aim is to increase the scientific relevance of legal research?[3]

2. THE ADVANTAGE OF LAGGING BEHIND

In other social sciences, quantitative research evaluation methods have been on the rise for the last couple of decades. Rankings of journals and publishers, h-indexes, journal impact factors, rejection rates, download frequencies and so on have become increasingly important. Especially policy makers, university managers, and funding agencies seem to have a preference for bibliometric indicators because these appear to be relatively objective, allowing them to compare the research output of scholars across national and disciplinary boundaries. Peer review, on the other hand, is often considered subjective, time-consuming and costly. At the same time, there is increasing criticism in

---

[1] Weiler, J (2012), 'Impact Factor – The Food is Bad and What's More There is Not Enough of It' 23 (3) *EJIL* 607.

[2] RAJ van Gestel (2015), 'Sense and non-sense of a European ranking of law schools and law journals', *Legal Studies*, Volume 35, Issue 1, pp 165–185.

[3] For a forceful critique of the marketization of universities more generally, see S Collini, *Speaking of Universities*, London: Verso 2017.

other scientific disciplines towards quantitative research evaluation methods, such as rankings and citation counts. Binswanger summarizes his concerns against what he calls 'governing by numbers' as follows:

> In the current system, scientific knowledge is replaced by measurable outputs. Not the content of an article or a project counts, but the number of published and cited articles or the number and the amount of money of the acquired projects. Since the measurable output is considered to be the indicator of quality, the true quality is more and more crowded out. The need to publish constantly leaves no time to worry too long about the progress of knowledge, although this should be the real purpose of scientific activity.[4]

This outcry is, unfortunately, not the sound of a single frustrated scholar. The essence of Binswanger's criticism that existing research evaluation methods run the risk of 'crowding out quality by quantity' is gaining more and more support from various stakeholders. In the San Francisco Declaration on Research Assessment from 2012, for example, a number of journals and scientific organizations protested against the abuse of bibliometric research assessment methods.[5] One of the recommendations was not to use journal-based metrics, such as journal impact factors, as a surrogate measure for the quality of individual articles in assessing a scientist's contributions or in making hiring, promotion or funding decisions. The reason is simply that a journal impact factor says something about the average number of citations that publications in a certain journal receive. It tells us little about the quality of an individual article in that particular journal and even less about the abilities of the author. Not only may poor-quality articles sometimes end up in high-ranked journals − for example, due to affiliation bias (e.g. journal editors being more favourable towards scholars from famous universities because these attract readers)[6] − but articles may be cited for different reasons. Certainly, not all of these reasons have to do with the quality of the substance of an article (e.g. think of negative citations; people citing an article because they disagree with the content; citations to show courtesy; self-citations used to promote oneself; citations to show command of the literature in the field).

---

[4] M Binswanger, 'Excellence by Nonsense: The Competition for Publications in Modern Science' in S Bartling and S Friesike (eds), *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*, Heidelberg: Springer, 2014, p 66.

[5] San Francisco Declaration on Research Assessment, accessed July 2018 at www.ascb.org/dora/.

[6] See A Yoon (2013), 'Editorial Bias in Legal Academia', *Journal of Legal Analysis*, Volume 5, Issue 2, pp 309–338, accessed July 2018 at https://doi.org/10.1093/jla/lat005.

Similar statements can be found elsewhere. In April 2015, a team of experts in the field of research evaluation from Leiden University published a manifesto in the journal *Nature* concerning the fact that research evaluation is increasingly led by data rather than by judgement.[7] They formulated ten principles for sound research evaluation:

1. Quantitative evaluation should support qualitative, expert assessment.
2. Research performance should be measured against the research mission of the institution, group or researcher.
3. Excellence in locally relevant research should be protected.
4. Data collection and analytical processes should be kept open, transparent and simple.
5. Those evaluated should be allowed to verify the evaluation data and analysis.
6. Variations by field in terms of publication styles and citation practices should be accounted for.
7. The assessment of individual researchers should be based on a qualitative judgement of their portfolio.
8. Research evaluators should avoid misplaced concreteness and false precision.
9. Systemic effects on a particular discipline of assessments and indicators should be recognized.
10. Quality indicators should be scrutinized regularly and be updated when necessary.[8]

A study concerning "Research performances in the humanities and social sciences" shows similar problems and presents a range of suggestions to remedy those, including the direct involvement of scholars in designing and implementing research evaluation methods. The study warns that quantitative assessment methods may be used in addition to qualitative evaluations, but should never replace them.[9] An overview of the criticism against the current use of metrics can also be found in the consultation document 'Science 2.0:

---

[7] Text and explanation of the Leiden manifesto, accessed July 2018 at www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351.

[8] For a systematic literature review concerning the effects of evaluation systems, evaluation practices and metrics (mis)uses, see S de Rijcke, P Wouters, A Rushforth, T Franssen, D Hammarfelt (2016), 'Evaluation practices and effects of indicator use—a literature review', *Research Evaluation* 25 (2): pp 161–169.

[9] A Hagall, J Lanarès, A Moraion, J Bregy (2018), 'Research performance in the humanities and social sciences', Final report, Swiss Universities, accessed July 2018 at www.performances-recherche.ch/uploads/Abschlusspublikation_P-3_EN.pdf.

Science in Transition', issued by the European Commission.[10] This consultation shows that there is general acceptance of the fact that traditional metrics are inadequate and that alternative ways to monitor 'open science activities' are necessary. Existing traditional metrics − such as citation scores, impact factors and download rates − are seen by many as inadequate proxies for research quality. However, this does not mean that there is nothing wrong with the most common alternative to quantitative impact measurement, which is peer review.[11] A wide range of complaints are also levelled against this research assessment method. To summarize the most important:

- There are not enough qualified reviewers for the increasing number of academic publications.
- Reviews are all too often conducted in a sloppy and superficial way because the review activity itself is usually not rewarded as a separate essential activity in the scientific community.
- Studies show that peer review suffers from all sorts of biases (e.g. confirmation bias, affiliation bias, ideological bias).[12]
- The selection of reviewers, and the instructions on what reviewers are supposed to pay attention to, are often not transparent.

These are not just theoretical risks. Empirical research shows that reviewers often have great difficulties in filtering out the best publications. Moreover, the inter-subjectivity rate between reviewers is often low, meaning that reviewers heavily disagree about the quality of individual manuscripts. This suggests that they either apply different (implicit) quality standards to assess manuscripts or may have strong personal preferences for certain types of research or a particular style of writing. Other problems related to research evaluation in the social sciences could also be mentioned, such as the presence of so-called 'citation cartels' (academics agreeing to cite each other in order to lift the impact scores for themselves or for a certain journal in which they publish), and the lack of reproducibility of certain research findings. The latter has become a major issue in the field of social psychology due to cases of research fraud that were not discovered for years; which raises the question to what extent academics in this field are critical of each other's research methodologies.[13]

---

[10] See http://ec.europa.eu/research/consultations/science-2.0/science_2_0_final_report.pdf, accessed July 2018.

[11] For an overview of current debate on peer review see the "peer review debate" organized by *Nature*; accessed July 2018 at www.nature.com/nature/peerreview/debate/.

[12] See, for instance, D Schatz (2004), *Peer review: A Critical Inquiry*, Oxford: Rowman & Littlefield Publishers, ch 2.

[13] See, for example, http://science.sciencemag.org/content/349/6251/aac4716, accessed July 2018.

The conclusion that legal scholars may draw from all this is that research evaluation in other (social) sciences is similarly flawed, which implies that we cannot simply copy the assessment methods in other sciences and apply them to academic legal research. After all, in that case we would probably also transplant the existing problems into the field of law. A fellow scholar from the social sciences even advised us: 'Please do not do what we have done, but do as we have learned!' In other words, law as a discipline has the advantage of being able to learn from the mistakes with regard to research evaluation that have already been made in other sciences. In the humanities, for example, colleagues tried to establish European journal rankings, but this project failed because the initiators did not include the scholarly community in the process and did not take into account the particularities of the field compared to other disciplines.[14]

3. BUT WE ARE SO SPECIAL AS LAWYERS…

In his book *Rethinking the Law School*, Stolker observes that European legal scholars are rather dismissive of measuring research quality via citation counts, rankings and other quantitative evaluation methods.[15] At the same time, though, the scholarly legal community also seems sceptical about external (double-blind) peer review, since many law journals in Europe still rely on editorial decisions alone. Apart from that, there appears to be no consensus on quality indicators that should be taken into account by authors, reviewers and editors in assessing the quality of different kinds of academic legal publications. Even more surprisingly for outsiders, there is still disagreement in different countries about how to distinguish academic publications from professional publications. All this makes it very difficult to compare the quality of legal scholars and their publications across national borders and sometimes even within the same jurisdiction. The question, however, is how long legal scholars will manage to maintain this exceptional position by claiming that law is different from other disciplines.

This does not mean that academic legal research has no special features compared to, for example, research in the field of science, technology and medicine (STM). As part of the social sciences and humanities (SSH), there

[14] See www.theguardian.com/education/2011/jun/27/journals-index-angers-european-academics, accessed July 2018.

[15] C Stolker, *Rethinking the Law School*, Cambridge: Cambridge University Press 2014, p 245.

are certain fundamental differences which make existing quantitative evaluation methods troublesome. Van Leeuwen mentions five:

1. While in the STM domains, scientific knowledge become obsolete within three to four years, in the SSH domains this might take up to ten years or more, while books influence the field sometimes even for decades.
2. Scholars in the field of STM are on average more focused on international cooperation.
3. Research teams in the field of STM are usually larger. In law, much research is still conducted on an individual basis. This also affects co-authorships.
4. The diversity of publication forms in the field of SSH is much greater than in STM, where journal articles are the dominant format. Moreover, in the field of SSH there is much more communication with audiences outside the academic world. In law, non-scholars even contribute to the academic debate because practising lawyers often publish in the same venues as scholars.
5. The variety of languages in which scholars publish in the field of SSH is much higher than that in STM. This alone makes bibliometric evaluations extremely complicated. As for legal research, Lienhard et al additionally mention the pronounced segmentation of academic legal discourse (e.g. the professional and academic audience of legal research; scholars in various fields of law have their own debates; there are regional, national, European and global discourses) and the sometimes atypical methods of traditional legal scholars (e.g. normative argumentation methods shared between the judiciary and the scholarly legal community).[16]

These characteristics make Van Leeuwen believe that standard bibliometric techniques, such as citation counts in Web of Science and Scopus, should be avoided in this field.[17] Perhaps these techniques could work for a number of (English language) journals in a few countries (e.g. the United Kingdom and Ireland), but certainly not for Europe as a whole − at least not for the coming years. Interestingly, De Witte has shown that it would be difficult even for a field of law that is transnational by nature: namely, EU law. He argues there is still no single 'European' discourse on EU law because there is no *lingua franca*. English scholars writing about EU law often do not read French and

---

[16] A Lienhard, F Amschwand, E Hermann, *Forschungsevaluation in der Rechtswissenschaft: Ausgangslage, Entwicklungen und Ausblick, LeGes* 2/2013, pp 411ff, accessed July 2018 at www.performances-recherche.ch/uploads/Lienhard_Amschwand_Herrmann+LeGes+2013_2.pdf.

[17] T van Leeuwen (2013), 'Bibliometric research evaluations, Web of Science and the Social Sciences and Humanities: a problematic relationship?' *Bibliometric Praxis und Forschung*, Band 2, 2013, p 8-1/8-18.

German sources and vice versa.[18] In practice, this means that debates about EU law in different member states can sometimes be quite narrow, because scholars have a limited overview of what is written about certain aspects of EU law in other member states. At the same time, however, due to the establishment of English language law journals − especially in the field of European and international law − publications in English are clearly on the rise. In certain other fields of law (e.g. multidisciplinary 'law and . . .' fields), this also seems to be the case, but this also raises questions.

In the UK and US, the local market for English language articles might be large enough to pretend to be 'international' even when the research as such deals with English or American law. This provides scholars in these countries not only with a language advantage (English being their mother tongue), but also with greater opportunities to publish about local issues, since scholars in these countries may continue to write about national law while being read throughout the Anglo-Saxon world, whereas foreign scholars will need to do something extra (e.g. comparative law) to get published in these same journals.[19] However, difficulties with the application of traditional bibliometric indicators and language barriers alone cannot explain why there is no uniform policy for peer review or editorial review among law journals and legal publishers throughout Europe. Several possible explanations can be offered here:

- Especially in smaller jurisdictions (e.g., the Netherlands, Finland and Slovenia), the scholarly community might be so small and intimate that (double) blind peer review makes no sense, because colleagues would still recognize each other's work.
- Journals and publishers do not believe that peer review delivers better (e.g., more objective) results than editorial review, or are afraid of the bureaucracy, time and costs that come along with peer review.
- External peer review of academic publications implies a certain consensus on how to discern between academic and professional publications, but this consensus might be absent in certain countries.
- Even where there is consensus about what constitutes an academic publication, there is not necessarily a shared understanding of the quality criteria that reviewers need to apply to manuscripts and about how to apply these

[18] B de Witte (2013) 'European Union law: a unified academic discipline?' in B de Witte and A Vauchez (eds), *Lawyering Europe: European Law as a Transnational Social Field*, Oxford, Hart Publishing, pp 101–116.

[19] For more in general on the SSH disciplines, see Y Li and J Flowerdew (2009), 'International Engagement versus Local Commitment: Hong Kong Academics in the Humanities and Social Sciences Writing for Publication', *Journal of English for Academic Purposes*, Vol 8, No 4 2009, p 279 (2009).

> to different types of legal research (e.g. doctrinal, comparative, empirical) in order to avoid bias.[20]

We do not really know why there is no established culture of peer review for law journals and legal publishers. Even where universities do not need the assistance of journals or publishers to determine the standards for publications and the review procedure, such as in the case of PhD dissertations, there appears to be a lack of uniformity. Doubts may be cast as to how legitimate this is. Roux, for instance, has defended that:

> The very notion of a PhD in law also assumes that some kind of cross-disciplinary standard must be met. Even when annual review panels are composed entirely of legal academics and where examiners are themselves academic lawyers, the criteria they are expected to apply are framed in cross-disciplinary terms. In effect this means that, however intuitive, and however resistant to specification in social science terms, the criteria for good doctrinal PhD research must be recast in terms that fit, or at least are not antithetical to, the standard elements of a sound doctoral research dissertation, viz.: a clear and confined research question, a comprehensive and targeted literature review, an appropriate methodology that is rationally related to a governing theoretical framework, and a plausible statement of the project's research significance. Two features of doctrinal research in particular make these standard elements difficult to apply: (1) the fact that doctrinal research is often presented in a highly rhetorical style; and (2) the fact that the criteria for sound doctrinal research as a whole, quite apart from doctrinal PhD research, are rarely articulated.[21]

That traditional legal research is often presented in a highly rhetorical style without articulation of the quality criteria that will be applied may be true; but the relevant question here is: to what extent is this justified by the special nature of legal research? Lovitts studied the quality indicators for PhD dissertations across ten disciplines.[22] She found out that there are remarkable similarities as to what scholars perceive as outstanding, average and poor dissertations. Criteria such as originality, methodological rigour, thoroughness, clarity and so on can be found almost everywhere. This does not mean that the way in which these criteria are implemented may not differ between disciplines. One of the peculiarities of the field of law, for example, is that the methodological

---

[20] See RAJ van Gestel (2017), 'Ranking, Peer Review, Bibliometrics and Alternative Ways to Improve the Quality of Doctrinal Legal Scholarship' in RAJ van Gestel, HW Micklitz and E Rubin (eds) (2017), *Rethinking Legal Scholarship: A Transatlantic Dialogue*, Cambridge: Cambridge University Press, pp 351–398.

[21] T Roux (2014), 'Judging the quality of legal research: A qualified response to the demand for greater methodological rigour', *Legal Education Review*, Vol 24, No 1, 2014: 173–200 at 181.

[22] B Lovitts (2007), *Making the Implicit Explicit: Creating Performance Expectations for the Dissertation*, Sterling: Stylus Publishing.

ground rules of the discipline are rather implicit, as Roux has argued with respect to PhD dissertations. Legal scholars are not used to thinking in terms of research design, hypothesis and theoretical frameworks. Nonetheless, this does not imply that it is impossible to develop more specific standards and methods for (the assessment of) academic legal publications.

Legal scholars perhaps need to realize that there is an important downside to keeping the quality standards for publications implicit. How can relative outsiders (e.g., potential employers) determine what a PhD title in law is worth if every university sets its own standards? How can legal researchers identify the journals to which they should submit an article if the standards for review are not transparent; how can they know what the best national and international journals are without some sort of benchmark? With respect to research grants, how should a funding body value the CV of a candidate who is applying for funding if there is no classification of journals or publishers or an alternative method to compare CVs? More generally, how can law schools, law journals and legal publishers show external accountability with respect to the quality of legal scholarship if there is no internal consensus about the quality standards for academic legal publications?

4. QUALITY OF LEGAL SCHOLARSHIP: A CONTESTED CONCEPT

One of the reasons why it might be difficult to develop more specific quality standards for academic legal publications is that quality is such a multi-faceted and relative concept. When deciding on the quality of an academic paper, for example, we probably expect more from a professor than from a PhD student; we would apply higher methodological standards for a monograph than for a case note; and most people would (hopefully) set higher standards for the English writing skills of a native speaker than for a foreign scholar for whom English is a second language. Things become even more difficult once we start asking ourselves to what extent we can apply the same standards to different types of academic legal publications: doctrinal, comparative, empirical and so on. Does methodological rigour, for example, not mean different things for various kinds of research? Replication of research results may, for example, be an important requirement for empirical legal research, but it is unrealistic to demand this with respect to doctrinal legal publications because, as Hutchinson and Duncan have argued:

> Many aspects of the law are contingent on context, and need to be interpreted and analysed for meaning. Synthesising the law and, where necessary, applying the law to the facts and context is a highly subjective process. Therefore the analytical, legal reasoning aspect of the process is necessarily a qualitative one. The outcome varies

> according to the expertise of the individual scholar and cannot be replicated exactly by another researcher. When a researcher undertakes doctrinal work, the outcome is totally dependent on the voice and experience of the individual. Doctrinal research requires a specific language, extensive knowledge and a specific set of skills involving precise judgment, detailed description, depth of thought and accuracy.[23]

If we cannot assess all legal publications according to exactly the same standards, how do we take relevant differences in individual publications into account? As far as the level of aspiration is concerned, there may also be differences between the quality of scholarly legal publications published in different outlets. A specialized law journal with a few submissions every month in a smaller jurisdiction cannot afford to set the same standards of excellence as an English language general interest law journal in the United Kingdom, with hundreds of submissions during the same period and with the same amount of issues published each year. From the perspective of potential authors, we can make a similar argument if we compare fields of law that are more transnational by nature, such as European and international law. After all, is it not far more difficult to get into the top English language journals in the field of international law, which receive submissions from all over the world, compared to writing about local law in non-English language journals that are read only in a single country? In other words, does it not make a difference whether journals or publishers (are forced to) set minimum standards as a threshold for discerning between sub-standard and acceptable publications, or can afford to apply standards of excellence to select only the best manuscripts for publication?

Things become even more complicated once we start to see that quality presupposes a standard to measure things of a similar kind for a certain purpose. Not only may doubts exist about the extent to which the same quality criteria apply to different kinds of legal research (e.g. doctrinal, comparative, empirical), but there also appears to be disagreement in Europe as to the purpose of legal research. Should scholarly legal research still be seen as a service to legal practice or has knowledge production become a goal in itself in the field of law? With regard to this question, a simple yes or no answer does not seem to be in order. Legal scholars face a dilemma: if they stay too close to practice, colleagues from other disciplines will not take them seriously; but if they distance themselves too much from legal practice, they run the risk of becoming irrelevant for those who make the law. In such a situation, it is not so easy to determine what the most important quality indicators for academic legal research should be. People simply have different expectations about academic

---

[23] T Hutchinson and N Duncan (2012), 'Defining and describing what we do: Doctrinal legal research', *Deakin Law Review*, Vol 17 No 1, p 116.

legal research. This makes the development of evaluation criteria all the more complicated.

5. REASONS FOR ACTION

In the introduction, we already mentioned that law as a discipline will be vulnerable towards criticism from funding bodies, government agencies and external evaluators as long as we cannot explain what quality standards are applicable for academic legal publications and how we monitor and evaluate our own research. It would be a serious mistake, though, to view the need for quality management primarily as a strategic answer to an outside threat. Law is a discipline in transition, in which competition, diversity, multi- and inter-disciplinarity, internationalization and methodological accountability are becoming increasingly important. We believe that the internal drive to respond to these changes should prevail over external pressure to change. If legal scholars throughout Europe wish to guarantee equal opportunities to be published and be recognized internationally, language barriers and differences in style of publishing somehow need to be overcome. The German *Wissenschaftsrat* formulated this for scholars in that country as follows:

> In order to be appreciated internationally in proportion to its true academic weight, German legal scholarship should closely follow and actively participate in European and international academic debates as well as in processes of making and developing the law. In order to do so, legal scholars should publish more frequently in foreign journals and integrate foreign literature into their own disciplinary discourses. This does not mean that scholars should merely shift the emphasis from German to English and begin publishing exclusively in English. But since legal scholarship is concerned with an object of inquiry that is constituted by language, the discipline should broaden its perspective and adopt a multilingual approach.[24]

Journals and publishers will therefore need to work together more transnationally in the future − for example, by translating manuscripts (for academic books, this already happens quite frequently) or reserving space for English language papers or summaries in more than one language.[25] If they do not undertake action in this respect, the chances are that more and more universities, individual scholars and funding bodies will publish their publications in open access repositories. Even in that case, it seems inevitable that publications will become more comparable, in terms of both format and evaluation

[24] *Wissenschaftsrat* (2012), 'Prospects of Legal Scholarship', Hamburg, p 15, accessed July 2018 at www.wissenschaftsrat.de/download/archiv/2558-12_engl.pdf.

[25] The latter is done, for example, by the Belgium journal for private law, *Tijdschrift voor Privaatrecht* (www.tpr.be/, accessed July 2018).

criteria, in order to offer a level playing field for legal scholars who want to address a broader European or global audience. Something similar applies to transnational mobility of academic legal researchers. After all, a precondition for hiring foreign scholars to conduct research is that universities can compare their academic abilities, of which publications are an important part.

Finally, multi- and inter-disciplinarity and the importance of methodological exposition (showing the relevance and validity of one's research) are important drivers behind the growing need to make our quality criteria and evaluation methods more explicit and transparent. Legal scholars nowadays increasingly study the law in its socio-economic, political and cultural context, and want to make recommendations for 'better law'. However, we often do not appear to realize that this implies that legal research is crossing boundaries with other disciplines and other research methods. Even for the most traditional form of legal scholarship, legal dogmatics, there is much to be gained from showing how the perspective of the legal scholar differs from that of the judge, the advocate, the legislator and so on.

To provide a simple example. For an advocate, it is perfectly fine to take the client's perspective and try to strengthen the arguments in its favour, while downplaying the counter-evidence and counter-arguments, because there is always an opposing litigating party. In scholarship, however, there is no 'natural enemy' for legal scholars; at best, there is a critical reviewer.[26] Such a reviewer usually studies a scholarly publication in isolation, which makes it complicated to identify suspicious patterns in the publication behaviour of an individual scholar. Furthermore, reviewers are often not focused on the possibility of bias or fraud. Empirical research by Chilton and Posner has shown,[27] however, how real and present the danger of biased legal scholarship might be. They studied a large sample of publications by constitutional law professors who donated money to the campaigns of either Democratic or Republican presidents and found out that those scholars who sponsored Republicans tend to lean more towards conservative viewpoints in the field of constitutional law, whereas the opposite was true for sponsors of Democratic presidents; while almost all constitutionalists claim to be objective. This might be dismissed as an anomaly by many European legal scholars, but how certain are we that the risk of biased and methodologically flawed legal research in Europe is so much smaller?[28]

---

[26] See R Spitzer (2008), *Saving the Constitution from Lawyers*, Cambridge: Cambridge University Press, ch 1.

[27] A Chilton and E Posner (2015), 'An Empirical Study of Political Bias in Legal Scholarship', 44 *J Legal Stud* 277.

[28] See, for instance, F Coomans, F Grünfeld and M Kamminga (2010), 'Methods of Human Rights Research: A Primer', *Human Rights Quarterly* 2010, Vol. 32, 179–186

Even though complying with methodological ground rules will not turn a mediocre publication into an excellent one, we believe that there is a relationship between methodology and quality. Methodological accountability, by showing how one has arrived (step by step) from a research question to a conclusion, can help to filter out flawed research. This can only function, though, when these ground rules are clear for everybody and when journal publishers, editors and reviewers know how to apply them. This could be another reason for making our implicit quality criteria and evaluation methods more explicit.

6. AIM AND RESEARCH QUESTIONS

The purpose of this book is not to take sides in the debate about the nature of law as an academic discipline (e.g. is it a science or not?), or to promote a certain view on quality management in academia. We feel that it is important to start with an observation of how academic legal research is currently being evaluated throughout Europe and of the criteria, indicators and assessment methods that are being applied. We did not start this project from a perfectly clean sheet, however. This project has a history informed by the desire of the Swiss rectors' conference, mentioned in the preface, to learn more about how scholarly legal research should be evaluated, according to both legal scholars and legal practitioners. This resulted in a first phase in which researchers from the universities of Tilburg and Leiden and Bern and Geneva undertook surveys[29] which showed, among other things, that legal scholars have a strong preference for qualitative evaluation methods, while law school managers

---

at p 182: 'human rights scholars tend to passionately believe that human rights are positive. Many of the scholars are activists or former activists in the field of human rights. Although seldom stated, the explicit aim of their research is to contribute to improved respect for human rights standards. They therefore risk ignoring the fact that the pursuit of human rights is not a goal in itself, but is merely one instrument designed to help improve respect for human dignity. They may forget that human rights standards are the result of compromises reached by states and may therefore be less than perfect. They may also overlook the fact that the mere adoption of resolutions by international bodies and the establishment of a new international institution will not necessarily result in the improvement of human rights on the ground'.

[29] For the results of the Dutch survey, see W van Boom and RAJ van Gestel (2017) 'Evaluating the Quality of Dutch Academic Legal Publications: Results from a Survey', *Utrecht Law Review* 13(3), pp 9–27. The results of the Swiss survey are published in A Lienhard, T Tanquerel, A Flückiger, F Amschwand, K Byland, E Hermann, *Forschungsevaluation in der Rechtswissenschaft: Grundlagen und empirische Analyse in der Schweiz*, Bern: Stämpfli Verlag 2016. A comparative analysis is published as: R van Gestel, K Byland, A Lienhard, 'Evaluation of Legal Research: Comparison of the Outcomes of a Swiss and Dutch National Survey', *Tilburg Law Review*, 2018/23(3), https://tilburglawreview.com/articles/6/.

appear to be more in favour of quantitative evaluation methods. Another result of these surveys was that the quality criteria that scholars and practitioners would like to see applied are still rather vague and abstract, and that participants in the project did not have an overview of the evaluation practice in their countries as compared to that in other countries. This was the main reason for writing this book: to '[g]ain an overview of the legal and policy norms, methods and criteria applied in the evaluation of academic legal research from a comparative perspective and find out where there is consensus and disagreement about current quality management practices that may affect legal scholarship'.

An additional reason for conducting the study was to see to what extent evaluation practices in the field of law are showing signs of convergence across countries. Apart from that, the country reports and the meeting with experts (see further below) offered a unique opportunity to share ideas and experiences with regard to how academic legal research is currently being evaluated throughout Europe. Finally, we wanted to learn whether further research could be useful in order to fuel a European-wide debate about the (assessment of) the quality of academic legal research. In order to achieve this aim, we formulated a number of relevant questions:

- Why are academic legal publications being evaluated in our sample of countries (what purposes does research evaluation serve)?
- How are academic legal publications being evaluated: which methods and techniques are being applied by the different evaluators (journals, publishers, funding bodies, research assessment exercises)?
- What sorts of consequences are attached to the outcomes of evaluations?
- To what extent is there a debate about the current or future course of action with regard to the evaluation of academic legal publications?

From the start, we realized that these questions should be indicative, because perhaps not every question is relevant for every country (e.g. for countries where there is no national research assessment, it might not make sense to discuss the evaluation of law schools and legal research centres in depth). Since the purpose of the research was to conduct an explorative study, we first wanted to gather factual information instead of opinions. Therefore, we developed a questionnaire on the basis of the existing literature (which is limited) and the Swiss and Dutch survey of legal scholars about existing evaluation practices. The initial questionnaire went beyond the scope of scholarly legal publications and included information with regard to the evaluation of research projects, legal researchers (e.g. tenure and promotion) and faculties and research institutions. In the end, however, this scope proved to be too broad for a book project, while not all categories delivered as much relevant information as we hoped for (e.g. some countries have no national assessment

of law schools or legal research centres). That is why we decided to delineate the scope of our research to academic legal publications.

7. METHODOLOGY

Since we learned from the first phase of the Swiss project on research evaluation that the body of literature on research evaluation in the field of law is fairly limited and mostly focused on particular countries instead of being comparative, we opted for a questionnaire, for which we developed a standard format (see Appendix I) in which four evaluation situations were identified:

- evaluations of academic legal publications (e.g. articles, PhD dissertations, papers);
- evaluation of legal research projects (*ex ante* or *ex post*);
- evaluation of legal researchers (e.g. tenure, promotion); and
- evaluations of law faculties and/or other research institutions.

We decided to select ten countries to include in our study. France, Germany and the United Kingdom were selected because these are the three dominant countries in the European Union, with strong academic legal communities. We wanted to include some 'old' and 'new' member states; hence, the Netherlands and Italy were selected as old member states and Austria and Slovenia as new member states. In order to strike a balance between northern and southern countries, we added Sweden and Finland and included Spain. Sweden and Finland are relative small countries in terms of their academic legal research communities, whereas Spain is a relatively big country with a sizeable legal research community. Moreover, with Switzerland, we also had a non-EU member state on board.

For the selection of rapporteurs, we trusted not only on our personal networks, but also the literature about legal scholarship in Europe. Daithí Mac Síthigh, for example, has previously written about legal methods in different law schools throughout Europe. As former research director, he was also highly familiar with the Research Excellence Framework in the United Kingdom. Ginevra Peruginelli did a previous study on evaluation practices in Italy. Elisabeth Maier and Marnix Snel wrote dissertations on the evaluation of legal research in Austria and the Netherlands, respectively; Marnix kindly took on the daunting task of writing a chapter about the EU's impact on research evaluation in member states. Antonina Bakardijeva-Englbrekt is not only a specialist in comparative law and familiar with the research culture in Eastern European countries, but has also worked for many years in Sweden at Stockholm University. Pia Letto-Vanamo was approached because she knows everything there is to know about Finland's legal research tradition and about the legal culture of the Nordic countries more generally. Janja Hoinik from

Maribor University came to our attention for her criticisms of recent developments with regard to the research evaluation system in Slovenia. As vice dean research, she also has a strong professional interest in our project. Delphine Costa had participated in a previous conference, organized by Andreas Lienhard and others, about research evaluation in Europe, where she shared her views on developments in France; we knew that she is also familiar with the Italian legal research community. Albert Ruda Gonzales kindly stepped in to write a chapter on Spain after we received a first draft written by Spanish experts in the field of research evaluation who acknowledged that they knew too little about the particularities of academic legal publishing in the country. Kai Purnhagen and Niels Petersen also stepped in at the last minute to write a chapter about Germany after we asked them to review a draft from another German colleague, which turned out to be too thin. Special thanks are due to Albert, Kai and Niels for helping us out at such a late stage of the project, which enabled us to complete the book without losing important jurisdictions. Andreas Lienhard, Karin Byland and Martin Schmied were, of course, involved as project leaders and architects of the entire initiative; and Rob van Gestel came on board because he has previously written about research evaluation in the Netherlands from a comparative perspective.

The questionnaire was completed at an expert meeting in Bern on 17 February 2017. During this meeting, every expert presented the findings of his or her study, and after every two presentations there was time for questions and debate. Since the draft outcomes of the questionnaire were distributed before the conference, there was plenty of food for discussion. During the afternoon session, the editors of this book tried to look for common trends in the national reports. For this purpose, we developed a number of propositions drawn from the country reports that were brought into the debate. This part of the expert meeting was recorded and minutes of the meeting were distributed afterwards, with the opportunity to provide comments and criticism. During the closing session of the expert meeting, we decided that in reporting our findings, we should not simply follow the structure of the questionnaires. Not all topics appeared to be relevant for all countries (e.g. in some countries there is no national research assessment); and we discovered that the quality of different types of publications is already such a rich topic that we decided to focus on the evaluation of academic legal publications in the four contexts mentioned above (research assessments of institutions, journals and publishers, funding bodies and tenure and promotions). The experts agreed with this focus for the book, although they realized that it might require some additional research.

We provided the experts with a sample chapter after the meeting in Bern, in order to guarantee a minimum level of uniformity. Moreover, we had two rounds of editorial comments in order to fine-tune the chapters: the first in May-June after the draft chapters were submitted and the second in December

based on the semi-final chapters. We also circulated the introductory chapter and the conclusions to the experts for comments, corrections and adjustments before submitting the manuscript to the publisher.

8. SCIENTIFIC AND SOCIETAL RELEVANCE

As far as scientific relevance is concerned, we believe this to be the first book ever to compare quality criteria and research evaluation methods in the field of law in Europe from a comparative perspective, written by legal scholars. Current studies are usually focused on one particular topic, such as journal rankings or peer review practices, and concentrated on the situation in a specific country. These studies are often undertaken by experts in the field of research evaluation or bibliometrics and written in the language of the respective countries. More generally, there is not a very vivid discourse on legal scholarship issues in Europe, as there is in the US. This probably has to do with the fact that law is still a discipline primarily organized around national legal systems, although this is changing quite rapidly.

The importance of this study from a theoretical perspective is to shed more light on what legal scholars view as distinctive for the quality of academic legal research. As we have previously seen in the US, there are doubts in some European countries regarding the future of doctrinal legal research. Some view this type of research as too inward-looking, focused on technicalities and lacking theoretical relevance; while others simply see traditional legal research as a distinct type of research that is neither good nor bad in itself. Instead of copying a less fruitful debate about the advantages of multidisciplinary 'law and . . .' research over legal dogmatics in the US, we believe it is more interesting to ask ourselves how we could recognize and evaluate the quality of different kinds of legal research − doctrinal, comparative, empirical and so on. Learning how scholarly communities in different European countries deal with this is a logical first step. The same applies, *mutatis mutandis*, to different types of publications. Instead of engaging in a debate on the benefits of case notes or textbooks, we think it makes more sense to see whether different criteria and/or assessment methods might be needed to assess the academic value of different types of legal publications that are currently (still) being written by academics. Finally, we observed that there is disagreement in several countries with regard to the distinction between academic and professional publications; hence, we would like to learn more how different evaluators in different countries deal with this.

At a time when competition for funding and accountability have become so important, it is inevitable that legal scholars must engage in a debate about quality management. Today's legal scholars can no longer hide behind local practices, craftsmanship and implicit traditions, because the internationaliza-

tion of our research will affect everybody sooner or later. It would be foolish to rely on outsiders, such as policy makers and university managers, to determine for us how to decide what good legal scholarship is. As we have mentioned above, experience in other disciplines has shown how dangerous it can be to leave quality management to outsiders. After all, it is no coincidence that in fields where new public management has been introduced in higher education, we have seen lots of adverse effects, from 'salami-slicing' publications in order to increase research production to self-plagiarism and the formation of citation cartels to 'beat the system'. Since law does not have a tradition of research evaluation and quality management, we have been relatively lucky so far; but the outcomes of the country reports reveal that it is highly unlikely that law schools, law journals and legal publishers will be able to escape the quest for more accountability in this respect. If the scholarly legal community fails to take the initiative, it runs the risk that others will try to impose their quality standards upon it. Apart from that, we believe that a self-conscious and mature discipline should be capable of explaining and showing to the outside world which quality standards apply to different kinds of research and what sort of evaluation methods are used to verify this.