

2009

The evaluation of legal science. The V.I.R.-model for integral quality assessment of research in law : what next ?

Serge Gutwirth

Published chapter : draft version for academic use only
Please do only refer to the published version

GUTWIRTH, S.,

The evaluation of legal science. The VI.I.R.-model for integral quality assessment of research in law : what next ?

in EISENDRATH, H. & VAN BENDEGEM J.P. (eds.) *It takes two to do science. The puzzling interactions between science and society*, Brussels, VUBPRESS, 2009, 69-80¹

I. Every science has its own truth regime

1. Although many may have their doubts, legal science does exist, not only because law as an academic discipline should in any case be scientific, but also – and especially – because apart from legal practice there is an activity that aims to develop a corpus of reliable and sound knowledge of law. Needless to say, legal science differs greatly from experimental sciences, field sciences and social sciences, which are respectively built upon the setting up of laboratory experiments, *in situ*, or field research, or quantitative/statistical and qualitative methods. The requirements that have to be met for the developed knowledge to be valid differ from science to science: every science has its own truth regime within which and with which its practitioners have to work if they want to be successful and if they want to be acknowledged by their peers. What is true in physics, sociology or legal science will depend on a different set of requirements for each of these disciplines. Nevertheless, they can all be referred to by the generic term “science”, because they all aim at the collective and mutually tested development of robust and reliable knowledge by the practitioners of a science.

2. Thus the model of experimental science cannot serve as the model for other contemporary sciences, if only because the latter usually cannot isolate their object in an experimental space that can be precisely controlled and set up anywhere in the world. Other sciences do not validate the same type of claims, but this does not make less scientific because. A statistical sociological analysis of voting behaviour is indeed very different from the experimental demonstration of gravity. In other words, the validity and success of a scientific claim is territorial and cannot be captured by generalising (epistemological) theories about Truth, Objectivity or Method, but results from the particular collective processes and practices of a given science or discipline.

II. So also legal science

3. The legislative branch takes political decisions and converts them into legislation, which in turn provides the framework for the legal practice or the ‘law in action’. This legislative activity of representative bodies essentially belongs to the realm of politics, which is nicely

¹ This contribution is an adapted and updated translation of S. Gutwirth ‘Evaluatie rechtswetenschappelijk onderzoek. VI.I.R.-model voor integrale kwaliteitsevaluatie van het onderzoek in de rechtswetenschappen’, *Nieuw Juridisch Weekblad*, 2007, nr. 168, 674 - 678. The text has been translated and corrected by Wim Van Verre and Michael Whitburn (ITO, VUB).

expressed by the existence of a special *section législation* in the Belgian *Conseil d'Etat* that watches over the legal quality of the process in order to optimise the conversion from political decisions into the statutes that provide the framework for the legal practice.

The legal practice occurs primarily in jurisdiction, in the courts. The law as such is *not* a science : unlike the sciences, it does not aim at the collective production of a body of robust and reliable knowledge, but tends to hold persons and things together in a web of relations that makes it possible to attribute responsibilities and impute acts, words and things to persons. It holds our societies together, via small and superficial, but *crucial* bonds, which put an end to conflicts and discussions, and stabilise relations, generating legal certainty.²

The production of legal science is the work of the legal doctrine, which endeavours to compile, describe and interpret aspects of law exhaustively, logically and systematically. It does so through monographies, articles, contributions to collections, reviews and not least through manuals that are constantly updated in accordance with the advancement of the law and that in turn, through teaching, wield influence over legal practice itself. Typically, a legal author will produce a systematising picture of important and pertinent juridical developments, sometimes followed by a retrospective analysis of these developments and their implications for future jurisdiction and/or legislation.

4. The legal doctrine shines structuring light upon the legal practice: it analyses the course of law and produces an ordered and systematic body of legal knowledge. Writings by legal authorities render the law teachable and transferable, and are an inspiration to legal practitioners who give life to the law. Thus, the legal doctrine occupies a remarkable and singular position: it is both the science of law, and an indirect, but often authoritative and influential source of that law. This explains the close ties between the legal science and the legal practice, the frequent combination of an academic career with a practical one, and it is also the reason why legal scholars address not only their peers in legal science, but must also be pertinent for legal practitioners. Legal science is therefore characterised by a double constraint : not only does it need to pass the quality test of fellow legal scientists – and must therefore be ‘robust’ as scientific knowledge – but it should also be relevant and pertinent to legal practice. Legal science has both scientific and legal dimensions that partly overlap and intertwine.

5. Apart from the more classic legal doctrine, legal science also encompasses comparative law and legal theory insofar as they maintain the same relationship of reciprocity with the legal practice and accordingly produce legally relevant knowledge. This distinguishes them from legal anthropology, legal ethnology and legal philosophy. As is the case for legal history, legal sociology, criminology and the science of public administration, these disciplines are not part of legal science since they do not generate legal knowledge, but another science’s knowledge about the law. These ‘metajuridica’ (‘about the law’) concentrate upon the study of the law from the perspective of sociology or psychology; as a result, they derive their scientific validity mainly from the truth regimes of these sciences. Legal science is autonomous: it includes legal doctrine, comparative law and legal theory insofar as these disciplines produce legal knowledge that is tested for resistance and pertinence by other legal scientists. Consequently, legal science has its own truth regime and should not be approached using criteria stemming from other disciplines or abstract methodological rules deemed to be relevant in all scientific disciplines.

² Cf. Br. Latour, *La fabrique du droit. Une ethnographie du Conseil d'Etat*, Paris, La Decouverte, 2002, 318p. See also : S. Gutwirth, P. De Hert & L. Desutter, ‘Technology regulation : the trouble with Lessig’s ‘optimal mix’ in R. Brownsword & K. Yeung, *Regulating Technologies*, Oxford, Hart Publishers, 2008 (forthcoming)

III. Measuring scientific output

6. In the context of the 'knowledge economy', the pressure on academics to raise their output has increased sharply. The opportunity to collectively explore and build up knowledge at one's own pace is gradually losing ground to a short-term achievement culture that emphasises the importance of tangible and visible results in terms of publications (*publish or perish*), doctoral degrees, secured research projects and subsidies, and patents (which indeed represent the perfect osmosis between knowledge and economy entailed by the knowledge economy). This shift has an impact not only on the assessment of a researcher's achievements, but also on the institutional policies of universities, for their funding is also increasingly dependent on scientific output. All this leads to a dire need for serious and reliable instruments for the virtually permanent evaluation and measurement of scientific output by researchers, research groups, fields of study, faculties, and academic institutions and higher education associations as a whole.

7. Many are rightly convinced that the best way to assess research and scientific publications is *peer review*: a thorough qualitative (and quantitative) judgement of a scientific result or file by a panel of both national and international researchers with expertise in the domain under investigation. However, considering the increase in the number and intensity of required evaluations, there is generally too little time and availability for this kind of ideal scientific assessment by peers. Moreover, policymakers demand figures because only figures can provide a basis for the allocation of funds. Finally, there is need for intercommensurability and comparability between the different sciences. As a result, quantitative measurement and registration of scientific output is seen as increasingly important. Whether one likes it or not, quantitative measuring instruments today dominate the scene, along with bibliometry and performance indicators that are said to correlate to the quality of the research as assessed by peers.

8. A number of scientific disciplines rely (to different extents) on the lists of science journals compiled by the *Thomson Institute for Scientific Information* (ISI), an American-based private enterprise. It selects the journals in accordance with a number of criteria among which, so the institute proclaims, the most decisive are frequency of issue, timeliness of publication, existence of peer review, use of English and the application of 'international editorial conventions'.³ Typically, a journal with a high ISI-score is written in English, is international and peer reviewed. Its multinational editorial staff presents an anonymised version of the submitted contributions to a number of anonymous referees who specialise in the matter discussed in the article. ISI also calculates the 'impact factors' of scientific journals and conducts author-related citation analyses (e.g. the 'h-index', the 'highly cited researcher' and more of the like). In general, the extent to which a scientific discipline relies on ISI depends partly on the context independence of the discipline in question (it is largest in experimental sciences), and partly on the 'coverage rate' of the periodicals analysed by ISI in that particular scientific field.

9. However, for many sciences, and for social sciences in particular, the ISI-system is irrelevant as it reports far too few effectively used publication channels (if only because these

³ The Thomson Scientific journal selection process, <http://scientific.thomson.com/free/essays/selectionofmaterial/journalselection/>, last consulted on 14 June 2008.

sciences are - also - practised in languages other than English). The system is problematic for other reasons as well: on the one hand, it forces these scientists to conform to requirements that are alien to or irreconcilable with their context dependent or local disciplines and on the other hand, it patronises non-ISI-disciplines. Needless to say, human sciences are much more closely related to the cultures, languages and societies in which they thrive and function. And so they should, if they are to be scientifically and societally relevant: a Belgian political scientist with a keen interest in Belgian politics can hardly publish only in English, especially if he is writing about theoretical implications of local politics that are not necessarily relevant to his English-speaking peers. Besides, why should his Dutch or French publications, by definition, be of inferior quality ?

10. Precisely because it is impossible to answer this question does the ongoing *European Reference Index for the Humanities* (ERIH)-project of the *European Science Foundation* (ESF) merit attention.⁴ Addressing the need to further ‘benchmarking and commensurability of quality’, this project attempts to draw up a list of science periodicals in 15 disciplines in the humanities (e.g. archaeology, art science, history, philosophy, psychology, anthropology ...). Unlike the ISI-system, ERIH does take into account both the European diversity of languages, cultures and intellectual traditions and the specificity of the disciplines in the humanities (such as the fact that the range of output channels is much broader than ‘scientific journals in English, and the fact that the life cycle of research is much longer here than in other sciences). More concretely, after consultation of the scientists involved, a list of peer reviewed scientific journals is compiled. Regardless of the language in which they are published the journals are ranked A, B or C, according to their circulation and reputation. The result is that Dutch, Portuguese or Hungarian scientific journals are not automatically eliminated, but are sometimes ranked B or even A. C-journals are typically of local or regional importance, and receive through this approach a weight or a value in a context of international comparability. The field of legal science is not involved in the project, which is most unfortunate, as it could offer perspectives at least as far as the output under the form of articles in scientific journals is concerned.

IV. Measuring the output of juridical science

11. The Belgian legal scientist who is confronted with ISI bibliometrical requirements and realises he is expected to take them seriously cannot but feel outraged for reasons closely connected to a number of immediately apparent features and practices specific to legal science. First, the Belgian legal expert publishes mainly in Dutch and in French, because they are the languages of his legal system, of his original legal sources, of the peers who deal with the same law in the same languages, and obviously of the law practitioners that he wishes to reach. So, law systems are strongly marked, locally ; they are highly sensitive to context and so differ greatly from place to place. To publish in English in an ISI journal is therefore an exceptional occurrence, all the more since most of the relatively small number of legal ISI journals are not really ‘international’, but mainly American or British. Moreover, as common law and continental legal systems are technically very different, it is difficult and sometimes impossible to correctly translate the French or Dutch Belgian legal concepts into English, since not all legal mechanisms have their equivalents in the other legal system.

⁴ *European Reference Index for the Humanities* (ERIH), *European Science Foundation* (ESF), via <http://www.esf.org/research-areas/humanities/activities/research-infrastructures.html>, last consulted on 14 June 2008.

There is also the fact that in legal science books and chapters in a book are at least as highly esteemed as, if not more highly than, articles in journals. And finally, legal journals are not only aimed at a readership of scientific peers, but more generally – and especially – at legal practitioners, which ties in with the particular nature of juridical science.

12. There is also the particular way in which legal research is carried out. Most research is not large-scale and networked, but it is often conducted by individual researchers working on an exhaustive, systematising and interpretative analysis of a legal subject. Legal scholars are text scientists, legal scientists are text scholars. And this also influences the way in which the legal doctoral thesis and legal project research are approached : the former is long drawn out, heavy, exhaustive and lonely hermeneutic journey through a sometimes very narrow subject ; the latter is simply scarce.

13. It is impossible to assess legal research on the basis of international bibliometric evaluation techniques more or less widely accepted in other scientific disciplines. To negatively evaluate a Belgian legal researcher because he has too few ISI publications is the same as saying that a Thai restaurant is no good because there are no chips or pizzas on the menu. If legal science has its own validity criteria it should also have its own evaluation criteria. Indeed, the best solution is to use the thorough quality-content assessment and comparative peer review ; but for reasons already explained, the peer review is no longer up to the task. There must also be more quantitative measurements. The question is : what is available ? Is there an instrument besides qualitative assessment by peers that can be used in Belgian legal research to measure output quantitatively ?

14. The answer is ‘no’. Yet the question is far from unimportant, because how else can a legal researcher claim that his scientific output is comparable to that of a physicist or sociologist ? Or how can he show that his research deserves to be publicly funded, or that his scientific output actually generates some of those funds ? Or, to put it more trivially, how can I convince a colleague in the immunology department without bringing a smile to his lips that the ‘comment’ I have just published on a judgment passed by a lower court is in fact a serious scientific publication ? How can he be convinced that some of the funding that he intends to devote to his AIDS research can just as usefully be spent on my legal research work into the legal aspects of autonomous agents ?

The answer to these questions is obvious : I need to be able to rely on a specific and transparent system of evaluation, supported by a network of peers / legal scientists, that allows me to align my own legal research work and scientific credentials with those of my legal colleagues. Should it appear that my work is the best in legal science and his is not the best in immunology, then he will have to accept that mine takes precedence over his. In other words, there is need for a clear, reliable and valid system of evaluation of legal research that meets the truth regime of the legal science.

V. The VI.I.R-project with relation to the evaluation of legal research

15. The need was recognised and addressed by VI.I.R (Vlaamse Interuniversitaire Raad), which produced a Model for Integral Quality Assessment of Research in Law in 2004, after two years of preparatory work carried out by an inter-university study group.⁵ Despite the (sometimes very convincing) basic reservations about the quantitative evaluation of research,

⁵ VI.I.R., *Model for Integral Quality Assessment of Research in Law*, 22 September 2004
<http://www.vlir.be/media/docs/Onderzoeksbeleid/notitieKZRechtenEngels.pdf>, last consulted on 14 June 2008

it has to be said that this model meets the specific requirements of juridical science : it attaches more value to books⁶ ; publications in Dutch published in national journals receive due recognition and bonuses are awarded for international publications ; it recognizes the fact that legal journals have a practical as well as a scientific orientation ; it adopts a cumulative approach in as much as more publications in lower ranking journals can compensate for there being no publications in the highest ranking journals. Evaluation systems focusing on the legal field were also developed with relation to project research, the coaching of doctoral research work, conferences and lectures, membership of editorial boards, etc.

16. The model as such has been accepted by the Flemish law faculties, and it was decided to put it to the test. The journals committee was set up and was given the task of evaluating Flemish legal journals on the basis of five issues provided by the journals' editorial boards. Each journal was awarded an A, B, or C-ranking. The Committee was set up on an inter-university basis and included four foreign, but Dutch speaking, members out of a total of nine. When the V.I.R. publicised the committee's ranking proposal on their website, there was a radical change in the goodwill attitude which the project had enjoyed until then.⁷ The ranking system became a subject of controversy in legal academic circles as well as in the related editorial boards of the journals considered, so that the project was soon deprived of oxygen and began to die a slow death. Having completed its task the publications committee was disbanded. But the books committee, which had been set up around the same time and was about to begin its work, was also disbanded, because of the commotion caused by the ranking system, and was replaced by a committee of exclusively foreign members. This committee never met and produced no work.

17. In March 2007, the chair of the V.I.R. asked the conference of deans of the Flemish university faculties to break the deadlock. It would appear that in a yet to be published document the conference of deans has recommended adopting the Dutch system to ensure that a reliable gauge may be developed 'rapidly and without undue effort'. Flemish law faculties should be consulted in the process. Meanwhile, however, the V.I.R. has launched a new initiative, spanning the whole field of social sciences and humanities (including law) to foster the development of a Flemish academic bibliography, which will officially (for financing purposes) list and rank the valuable scientific publications (not only journal articles). But in this project, the specificities of the legal discipline may again be disregarded by the bibliometrical customs of the other social sciences, which very often rely on criteria similar to those used by ISI ('international' journal publications in 'English' as 'top publications').

VI. The Dutch report 'Towards performance indicators for research in legal science'

18. The report 'Towards performance indicators for research in juridical science'⁸, which was completed in March 2007, followed the report 'Passing judgement on law' of 2005⁹. Both

⁶ In some disciplines the writing of books is seen as a negative sign: only scientists who are unable to have their writings published in ISI-journals would write books, which are deemed to be published for commercial reasons and without serious peer review. How wrong this assumption is, at least in the field of philosophy, has been nicely described in : J. De Mul, 'Publish and perish', *Tijdschrift voor filosofie*, 2005, 426-428

⁷ *Operationalisering van het model*,

http://www.vlir.be/02thema/04onderzoeksbeleid/Ranking/Vlir_update_Ranking_041206.pdf last consulted on 14 June 2008

⁸ *Naar prestatie-indicatoren voor rechtswetenschappelijk onderzoek. Rapport van de Commissie Prestatie-indicatoren en ranking, ingesteld door het Disciplineoverlegorgaan Rechtsgeleerdheid van de VSNU*, Maart 2007, 47 p. <http://www.vsnunl.nl/web/show/id=51717/langid=43>, last consulted 14 June 2008.

⁹ *Oordelen over rechten, Rapport van de Commissie Voorbereiding Onderzoeksbeoordeling Rechten*, oktober 2005, 77 p.,

reports were produced on the initiative of the consultation body for disciplinary matters of the Association of Universities in the Netherlands (the VSNU). The final aim is to establish criteria for evaluating legal research. And clearly, the two Dutch reports often draw their inspiration from the work carried out by the VI.I.R.

19. The March 2007 Dutch document seeks to define evaluation criteria that can be applied to the specific and particular nature of legal science : it is largely a matter of explicating on the basis of performance indicators criteria of prime quality research that are already implicitly being used in the academic community of legal experts. Nevertheless, many observers have insisted on the limitations of such indicators for juridical research, and as corollary on the prominent role that peer review must play as the prime and best method of evaluation.

20. As far as the ranking of journals is concerned, the Dutch report does not opt for the VI.I.R model in which the A, B, or C score is based on scrutiny of the content by a committee of peers. The Dutch report is extremely cautious and points to the many drawbacks and limitations of a system of classification including the fact that a journal's ranking system reflects the mean quality of the journal and says nothing about the article itself. Consequently, the Dutch report relinquishes the A, B, or C ranking system and opts for the division into two categories instead : A-journals that meet particular criteria and will therefore be considered 'scientific journals', and other journals.

There are three conditions to be met by A-journals : the journal mainly addresses a scientific forum and always strives to be original, far-reaching and thorough ; submitted articles are checked by members of the editorial board against a pre-existing format in which minimum criteria for an article have been determined ; editorial boards mainly include recognised experts in the field and - should there be a lack of experts - peer reviewers.

Following the proposal, the list of A-journals will be drawn up in cooperation with the editorial boards of the journals themselves and in accordance with a particular procedure in which the journals will be required to say whether they meet the pre-set conditions.

21. The Dutch approach is different from the VI.I.R model in as much as it uses much broader categories and is based on the active involvement of editorial boards throughout the process. As a result, there may be far less commotion and controversy, and it is quite possible that a great many A-journals will appear that will cover every sub-speciality. On the other hand, there is the question whether such an approach will be sufficiently discriminating and whether it will have enough credibility in other disciplines where there is a scarcity of A-journals. A similar system could be adopted for books, which could be regarded as scientific provided they are included as part of a series with a permanent editorial board, or provided they have previously been submitted to an independent peer review.

What is striking is the fact that the Dutch model shows no concern for preserving the anonymity of the peer review, and that articles can be evaluated by expert editorial boards. This is certainly in line with current editing practices in the legal field. The Dutch report also recommends that publications that appear in journals published in Flanders should take into account work carried out by the VI.I.R – which is a problem since the VI.I.R.-initiative has been aborted.

What now ?

The V.I.R model was one that could meet the new – possibly unsavoury, but nevertheless very real - challenges of a changing academic and scientific world, and one that could do so without losing face, while keeping an eye on the link between the quality of the manuscript and its evaluation. More so even, it represented a transparent response to the challenge of putting legal science on the academic maps alongside other disciplines. Seen from this angle it was a strong and courageous initiative that corresponded to a real and concrete need. Obviously, there was also room for improvement, which is why it was designed as a pilot scheme that was first meant to be tried out before it could be implemented.

The Dutch model sets out to achieve the same aims, but using other means and techniques. The legal experts concerned will – at least at first - probably find it less controversial, but it will be much less convincing for our colleagues in other scientific disciplines, because it is far less selective. If, say, half the current legal journals become ‘A-journals’, then the comparability – both internally and externally - of legal-scientific research will remain small. And wasn’t that precisely the problem that had to be solved ? How can I convince researchers from other disciplines that legal researchers are also engaged in serious scientific work, even though the work may be carried out in a very different way ?

Serge Gutwirth is a professor at the Law and Criminology Faculty of the Free University of Brussels. He was a member of the ‘journals committee’ of the V.I.R. project.