

WIBREN VAN DER BURG

The Merits of Law

An Argumentative Framework for Evaluative Judgements and Normative Recommendations in Legal Research¹

ABSTRACT: Is the law good? How can it be improved? These questions are frequently addressed, both in traditional doctrinal research and in interdisciplinary legal research. In this article, I elaborate a general argumentative framework for justifying evaluations and recommendations for legislative reform, and I identify the chains of argument for making evaluations and recommendations. This may help researchers to make their arguments explicit and transparent, and then to justify the choices made for each of the steps in the argument. This enables readers – as well as authors – to assess whether the argument’s conclusions are sound and convincing. A key problem in making these arguments concerns the choice of standards for evaluation. I suggest that they can be best understood in terms of the underlying values. In order to make all-things-considered evaluations and normative recommendations, interdisciplinary cooperation between doctrinal, socio-legal, and legal-philosophical researchers is necessary. Therefore, individual researchers will usually have to settle for pro tanto evaluations and recommendations.

Keywords: legal research methods, evaluation, recommendations, law reform, argumentative practice, chain of justification, argumentative framework, values.

Schlagworte: Methode rechtswissenschaftlicher Forschung, Evaluation, Empfehlungen, Reform des Rechts, argumentative Praxis, Rechtfertigungsstrang, Argumentationsrahmen, Werte

“Our inability to prove that we are right does not remove the need for justification. Nor does our desire to rest our normative claims on secure justifications protect us from the need to justify our actions and our laws even when it turns out that such foundations do not exist.”²

- 1 This paper has been presented at seminars at Queen Mary University London, Erasmus University Rotterdam, the University of Bristol, the University of Edinburgh, the Free University Amsterdam, and the IVR in Lisbon. My thanks go to members of the audiences for various helpful suggestions. I am also indebted to Peter Cane, Maks Del Mar, Donna Devine, Jeanne Gaakeer, Kate Malleson, Sanne Taekema, and Bart van Klink for helpful comments on earlier drafts of this paper, and to my student assistants Jacqueline Brand, Haris Sabanovic and Martine de Vries for their help in the literature research.
- 2 Joseph William Singer, Normative Methods for Lawyers, *UCLA Law Review* 56 (2009), 926

This material is under copyright. Any use outside of the narrow boundaries of copyright law is illegal and may be prosecuted.

This applies in particular to copies, translations, microfilming as well as storage and processing in electronic systems.

© Franz Steiner Verlag, Stuttgart 2019

1. Introduction

Is the law good?

How can it be improved?

Such questions arise frequently in the context of legal research. Law journal articles discuss, for example, whether a new statute fits in the system of law, whether there is tension or even conflict with fundamental legal principles, whether it will be effective, and whether it will promote the maximisation of wealth. Articles may also provide normative recommendations to legislators on how to improve the law, as well as suggestions on how the judiciary should decide future cases (especially in case notes).³

These evaluations and recommendations are often merely a short final section in a primarily doctrinal or socio-legal journal article. It is not uncommon to find articles in law reviews where 90 % of the text is devoted to a systematic presentation of the current law, with only a few paragraphs at the end devoted to evaluation and suggestions for reform. These evaluative and normative sections often resemble an attorney's brief in which the arguments lead directly to the conclusion.⁴ Counter-arguments and alternative solutions may hardly be mentioned and, if discussed at all, are easily dismissed. The standards of evaluation are often unarticulated and simply taken for granted. Moreover, many steps in the argument as well as many assumptions are not discussed explicitly but are left implicit.⁵

The main reason for this style is that most legal researchers have been trained as law students. In the Common Law tradition, law schools train their students as lawyers who can argue a case in an adversarial context. However, this also encourages a style

- 3 Sanne Taekema, *Relative Autonomy: A Characterization of the Discipline of Law*, in: Bart van Klink & Sanne Taekema, *Law and Method*, 2011, 35–36 argues that most doctrinal scholars go beyond mere reconstruction of the positive law, and include two additional aims: namely critical evaluations and recommendations for law reform. See also Carel Stolker, *Rethinking the Law School. Education, Research, Outreach and Governance*, 2014, 215. Terry Hutchinson, *The Doctrinal Method: Incorporating Interdisciplinary Methods in Reforming the Law*, *Erasmus Law Review* 8 (2015) 3, 130–138 refers to the Australian Pearce Committee that distinguished in 1987 between doctrinal and reform-oriented research. According to Hutchinson (this footnote, 132), however, the category of reform-oriented research was also a form of doctrinal research. “Most ‘good’ quality doctrinal research goes well beyond description, analysis, and critique, and invariably suggests ways the law could be amended or the philosophy, processes or administration of the law could be improved.” There may be a difference here between Civil Law and Common Law traditions. My impression is that in most Civil Law countries the overwhelming majority of legal research is still doctrinal, and basically focuses on exposition of the positive law, with evaluation and recommendation only as a side-product. In Common Law countries, the focus on expository work is less dominant, and straightforward normative research is much more common. Cf. this quote by a US scholar: Edward L. Rubin, *On Beyond Truth: A Theory for Evaluating Legal Scholarship*, *California Law Review* 80 (1992) 4, 903: “The purpose of legal scholarship is most accurately described as prescription or recommendation.” Few Civil Law scholars would subscribe to this sweeping statement.
- 4 See J. B. M. Vranken, *Mr. C. Asser's Handleiding tot de beoefening van het Nederlands Burgerlijk recht. Algemeen deel. **** Een synthese*, 2014, 105; Stolker (note 3), 208.
- 5 Rob van Gestel & Hans-Wolfgang Micklitz, *Why Methods Matter in European Legal Scholarship*, *European Law Journal* 20 (2014), 302 observe that all too often there is a ‘jump’ to normative recommendations, “without hardly any evidence that what is proposed will actually lead to the intended result.”

of reasoning that leads directly to the thesis students need to defend.⁶ Lawyers have to present a case in such a way that the audience is fully convinced, and that may involve a suggestive structuring of the narrative, the facts and the arguments, and the passing over of some of the counter-arguments. When reading a good case brief, one can hardly imagine that a reasonable person would disagree – until one reads the opponent’s brief. In the Civil Law tradition, students are trained to think more like a judge and to solve cases by applying ‘the law’. They learn to present positive law as a coherent doctrine on the basis of legislation and other legal sources. Until recently, writing an authoritative handbook presenting a field of law systematically was considered to be the highest accomplishment for legal scholars, and law students were mainly trained to write in this expository style.⁷

However, and as various authors have convincingly argued, while this style may be good for a practicing lawyer, it is not a good model for a legal researcher – not even for merely expository research.⁸ We may expect from academic researchers an attitude of openness, a careful presentation of all relevant considerations and options, and a balanced conclusion after an impartial weighing of all the arguments pro and con.⁹ Legal research requires the impartial analysis of all relevant facts and arguments, and this should be done not only in the research itself but also be presented explicitly in the publication so that the academic forum can determine whether the arguments are sound and convincing.¹⁰

Even legal expository research should acknowledge this open and discursive dimension, and instead of ignoring the tensions and inconsistencies in law we should be explicit in identifying them. *A fortiori*, a similar authoritative style is not acceptable in the context of evaluations and recommendations. The reader should be able to assess the arguments. In academic publications, choices should be made explicit and be justified, and the argument should consist of a balanced debate with regard to pros and cons. But how precisely should this be done?

6 Lee Epstein & Gary King, *The Rules of Inference*, *University of Chicago Law Review* 69:1 (2002), 9; Van Gestel & Micklitz (note 5), 303 with additional references.

7 An illustration is how many Dutch treatises or handbooks present only one interpretation of the law (that of the author), and then mention alternative views only in a footnote with the mere addition of *Anders*: (“Differently:”)

8 See Vranken (note 4), 105; Stolker (note 3), 208. Rob van Gestel, Hans-W. Micklitz & Miguel Poiars Maduro, *Methodology in the New Legal World*, 2012 (EUI working paper), 5 argue: “Especially the training of law students in the ‘art of persuasion’, which is an important part of the education of lawyers in many law schools, not only in the U. S. but to a lesser extent also in Europe, is tricky from a methodological perspective since it can easily spur one-sidedness and may provide an impetus to leave aside sources, arguments, and opinions that do not fit with the claims one wants to make.”

9 Of course, scholars also need to pay attention to the rhetorical dimension. Appealing to convincing cases and examples are part of that, but biased narratives and the selective use of materials should be avoided.

10 Of course, this implies that if law schools claim to provide an academic curriculum – and I believe they should – they should also teach different styles of reasoning and research other than just the familiar ones of legal practice.

This brings us to the question central to this article: *How can legal researchers, in a methodologically sound way, justify evaluative judgements concerning the merits of the law, and how can they justify normative recommendations about improving the law?*

The research question in turn raises a host of further questions. These include, for instance, what are the criteria for calling a law ‘good’, or, in a broader sense, for assessing the merits of a law? How can we make sound and convincing arguments for evaluations? How do we decide which possible solutions should be analysed? And how do we assess realistically which of these alternatives is best, all things considered? Obviously, one article cannot fully answer these questions. What it can do, however, is take the first step. To this end, I will develop a general argumentative framework that could serve as the starting point for future elaborations of research methods. The focus will be on the formal structure of the argument. I will identify the steps in the argument – the elements – but will not discuss how they can be substantiated.

This article proceeds as follows. In Section 2 and 3, I will position and refine the research question against the background of the recent debate on methods of legal research. In Section 4, I will analyse the characteristics of evaluative and normative arguments in the context of legal research. A crucial issue concerns the standards of evaluation, which will be discussed in Section 5. Next, I will elaborate a chain of argument for evaluation – the questions that must be addressed for a full and sound argument (Section 6). If we want to go from an evaluation to normative recommendations, we need to address a number of additional issues; this will require a further chain of argument, to be found in Section 7.

The framework I develop may at first glance seem way too ambitious. Therefore, an obvious objection to the approach sketched here could be that I am asking too much. This objection is discussed in Section 8. I will argue that the character of legal research – like law itself – is gradual. It is an enterprise or project in which we may succeed to a lesser or a greater extent. Therefore, even if the ideal of fully sound evaluations and recommendations can never be reached, it may still be helpful to develop such an ambitious framework as a regulative ideal (Section 9). This may enable us to attain a higher degree of realisation of the ideal. I will conclude with a section containing suggestions for future research. There are still many gaps in this general framework, and I will identify some of them and suggest how they might be addressed.

This article is not aimed simply at methodological purists. It is important that we explore ways to improve the quality of arguments that justify evaluations and recommendations. Improvement of the argumentative framework and the research methods will lead to a higher quality of legal research and to a better warrant for the evaluations and recommendations that result from it. Moreover, it will also help beginning researchers with their research design and with the presentation of their results. Obviously, these improvements are not merely in the interest of legal research itself. If, in current legal research, the evaluations and recommendations are methodologically weak, this means that recommendations for – or against – legal reform may sometimes be ill-advised, and even simply wrong. If legislators act on the basis of that advice, we may end up with a lesser quality of law.

2. Restricting and refining the research question

How might we address this complex issue? In the Introduction I have already indicated one important restriction. This article will focus on the formal structure of the argument, which means that I will try to be as neutral as possible here.¹¹ Ph.D. candidates and more experienced researchers of every normative and disciplinary approach should be able to use it as a framework. Obviously I have personal preferences for certain types of research and for certain normative positions, but I have restricted these to the footnotes.¹²

Further distinctions and restrictions are still necessary, however, to make the project feasible. First, I will confine myself to one specific class of recommendations: namely, those directed at changes in the substantive contents of the law. Legal research may also include other recommendations: for example, that the police should spend more time enforcing certain rules, that statutory norms should be more effectively publicised, and so on. However, because including such recommendations would make this an extremely broad project, I will focus on a distinct subclass of recommendations, which I will simply call *normative recommendations*, as they propose changing legal norms.¹³ This specific focus has been chosen because normative recommendations present a special challenge: they can only be fully justified by taking into account the internal perspective of law. An external perspective, such as that of law and economics or Rawlsian political philosophy, may be enough to justify a *pro tanto* evaluation of the law, a partial evalu-

- 11 Therefore, for the purposes of this article it does not matter whether a doctrinal researcher has positivist, interpretivist, or natural law leanings. The debate between legal positivists and their opponents may be – according to some interpretations – about the question of whether we can identify and reconstruct the contents of the law without reference to underlying moral values. In Austin's famous words, quoted by H. L. A. Hart, *Positivism and the Separation of Law and Morals*, *Harvard Law Review* 71 (1958), 596: "The existence of law is one thing; its merit or demerit is another." I do not think it is helpful to structure the debate between legal positivism and its opponents in the way Austin and Hart did it. Even natural law theorists will be able to distinguish both questions. See Wibren van der Burg, *The Dynamics of Law and Morality. A Pluralist Account of Legal Interactionism*, 2014, 65 f.; Dan Priel, *Toward Classical Legal Positivism*, *Virginia Law Review* 101 (2015), 989–991. Doctrinal researchers belonging to each of the competing traditions all claim to be able to reconstruct and describe the contents of the law in their own way. After that description each of them can always ask the separate question: what about the merits of the law thus understood?
- 12 In various publications, I have advocated interdisciplinary cooperation between doctrinal researchers, socio-legal researchers, and legal philosophers. See Sanne Taekema & Wibren van der Burg, *Towards a Fruitful Cooperation between Legal Philosophy, Legal Sociology and Doctrinal Research: How Legal Interactionism May Bridge Unproductive Oppositions*, in: R. Nobles and D. Schiff (eds.) *Law, Society and Community. Socio-Legal Essays in Honour of Roger Cotterrell*, 2014, 129–145 and Van der Burg (note 11, arguing for sociologically informed jurisprudence). This cooperation is a central theme in the work of Roger Cotterrell, e.g. Roger Cotterrell, *Law's Community. Legal Theory in Sociological Perspective*, 1995, Oxford: Clarendon Press and Roger Cotterrell, *Law, Culture and Society. Legal Ideas in the Mirror of Social Theory*, 2006. See Cotterrell (this footnote, 2006), 29: "Legal philosophy and legal sociology are co-workers in a common enterprise of legal explanation."
- 13 Of course, there is not always only one uniquely right solution to a problem. Recommendations could also suggest two or more alternative solutions to a problem; there need not always be conclusive reasons to prefer a specific solution. In a weak form, they could even simply imply the rejection of one or more possible solutions.

ation from a specific perspective. For normative recommendations, however, we need to take into account the distinct characteristics of law, not only of law in general but also the specific characteristics of concrete legal orders. For example, many bio-ethicists have argued that certain types of euthanasia can be morally justified. In order for them to go from this position to legal reform, they need to take into account legal characteristics such as that the law cannot easily assess intentions, that the law has specific problems regarding proof, and that laws have to be general and should not be tailored to exceptional cases, as well as the doctrinal understanding of *mens rea* in a specific jurisdiction. Thus the findings of interdisciplinary research must be translated back into legal terms and incorporated into the legal perspective. This problem of incorporation is a crucial challenge needing to be addressed.¹⁴

Second, this article will focus only on legislation. The general research question is, of course, also relevant with regard to evaluations of case law and recommendations to the judiciary on how to decide future cases. However, the differences between Common Law and Civil Law traditions are quite substantial with regard to the role of legal scholarship in the development of case law.¹⁵ These variations would complicate a more general analysis. For the purpose of this article, therefore, the focus will remain on statutory law.

Third, what is legal research? In a broad sense, it can be understood as all research on law. This includes doctrinal research and interdisciplinary research, as well as non-doctrinal monodisciplinary research such as economics, philosophy, history, or sociology. In this article, I will focus on doctrinal and socio-legal research, simply because they are the most common types of legal research.¹⁶ Both types can lead to evaluations and rec-

- 14 See the two special issues of the *Erasmus Law Review* on ‘The Incorporation Problem in Interdisciplinary Legal Research’, 2015, edited by Sanne Taekema and me. See also Frans L. Leeuw with Hans Schmeets, *Empirical Legal Research. A Guidance Book for Lawyers, Legislators and Regulators*, 2016, 220–235; Ivo Giesen, The Use and Incorporation of Extralegal Insights in Legal Reasoning, *Utrecht Law Review* 11 (2015) 1, 1–17. Péter Cserne, Making Sense of Nudge-Scepticism: Three Challenges to EU Law’s Learning from Behavioural Sciences, in: Alberto Alemanno, Anne-Lise Sibony (eds.) *Nudge and the Law: A European Perspective*, 2015, 281, makes an interesting distinction between “three kinds of challenges to the integration of empirical knowledge about human behaviour into (EU) law: epistemic (related to how policy-relevant empirical knowledge is generated), institutional (how this knowledge is channelled and transferred into legislation and other legal procedures) and normative (related to law’s commitment to counterfactual models of human behaviour and non-instrumental goals).”
- 15 William Twining, Ward Farnsworth, Stefan Vogenauer & Fernando Tesón, The Role of Academics in the Legal System, in: Peter Cane & Mark Tushnet (eds.), *The Oxford Handbook of Legal Studies*, 2003, 920–949
- 16 I will not discuss the methods for doctrinal or empirical research here, but will simply assume that the doctrinal reconstruction of the positive law or the empirical research has been executed, or will be executed, in a methodologically sound way according to the methodological standards of that discipline. The question I want to address here is the further one of how to use the results of these doctrinal or empirical studies for evaluations or recommendations. On methods of doctrinal legal research, see Terry Hutchinson, Developing legal research skills: Expanding the paradigm, *Melbourne University Law Review* 32 (2008), 1065–1095; Terry Hutchinson, Doctrinal research: Researching the jury, in: Dawn Watkins & Mandy Burton (eds.), *Research Methods in Law*, 2013, 7–33; Hutchinson (note 3); Mátyás Bódog, Legal Theory and Legal Doctrinal Scholarship, *Canadian Journal of Law and Jurisprudence* 23(2010) 2, 483–514; Mátyás Bódog, Legal Doctrinal Scholarship and Interdisciplinary Engagement, *Erasmus Law Review* 8 (2015) 3, 43–54; Mark van Hoecke (ed.), *Methodologies of Legal Research*, 2011; Watkins and Burton (this footnote). For socio-legal research, see Cotterrell (note 12, both); Reza Banakar & Max Travers, *An Introduction to Law and Social Theory*, 2002; Reza Banakar & Max Travers, *Theory and Method in Socio-Legal Research*, 2005; Jonathan Ver-

ommendations. However, it should be noted that there is a crucial difference between these two types with regard to normative recommendations. In order to formulate sound and concrete proposals for statutory change, we need to phrase them in terms that can be recognised by legislators, and we need to determine whether these formulations fit in the conceptual framework of the legal order. This requires a participant's or internal point of view. This is also the stance usually taken by doctrinal research.¹⁷ Socio-legal research generally adopts an external point of view, and does not focus on whether a text would fit in the legal system.¹⁸ Therefore, for this type of research, a number of additional steps are needed to go from evaluations to normative recommendations. For example, on the basis of the evaluative conclusion that citizens do not follow a law because its text is unclear or even incomprehensible to them, we can recommend that the text should be clarified, but in order to determine precisely how, we need to take into account the conceptual framework of the legal system as a whole. With regard to evaluations, however, purely external approaches are possible and legitimate.

It is necessary to clarify the concept of evaluation. Hence, we should distinguish between *pro tanto* and all-things-considered evaluations and recommendations. A *pro tanto* evaluation is a partial one, taking into account only certain dimensions (for example, whether a statute fits in the legal system and conforms to the rule of law) but ignoring others (for example, whether the law is effective). An all-things-considered evaluation takes into account all relevant dimensions on which to assess the law, and thus requires a much broader interdisciplinary orientation. The focus in this article is on a framework for all-things-considered evaluations, but, indirectly, it may also help us to better understand the justification for *pro tanto* evaluations.

'Evaluation' is interpreted here more broadly than is often done in, for example, evaluation studies. *Evaluation in the narrow sense* usually involves assessing empirically whether a law has realised its explicit purposes and other valuable aims, and whether there are no negative side-effects.¹⁹ In other words, it provides empirical data on a number of given evaluation criteria. However, in order to evaluate in the broad sense, we must also determine the importance of each of the criteria, and how a negative result with regard to one criterion is to be weighed against a positive result concerning another dimension. For this, we need a theory of values and their relative importance. It is like

schuuren (ed.), *The Impact of Legislation. A Critical Analysis of Ex Ante Evaluation*, 2009; Simon Halliday & Patrick Schmidt, *Conducting Law and Society Research: Reflections on Methods and Practices*, 2009; Fiona Cowney & Anthony Bradney, Socio-legal studies: A challenge to the doctrinal approach, in: Watkins & Burton (this footnote), 34–54; Leeuw & Schmeets (note 14).

17 Taekema (note 3) at 41: "The researcher takes part in the practice and reflects upon it in the same terms as the other participants. This does not mean that he needs to accept the practice uncritically, but it does mean that he needs to explain his criticisms in terms that the participants of the practice understand."

18 On the distinction between internal and external perspectives, see esp. Brian Tamanaha, *A Socio-Legal Methodology for the Internal/External Distinction: Jurisprudential Implications*, *Fordham Law Review* 75 (2006), 1255–1274.

19 Examples are Ray Pawson, *The Science Of Evaluation. A Realist Manifesto*, 2013; Michael Quinn Patton, *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 2015; G.J. Veerman (with R.J. Mulder & E. S. M. Meijnsing), *Een empathische wetgever. Meta-evaluatie van empirisch onderzoek naar de werking van wetten*, 2013.

grading a student paper: teachers have to assess whether the paper shows mastery of the literature, creativity, convincing arguments, good style, and so on, but they must also decide on the relative importance of each of these criteria. *Evaluation in the broad sense* involves both assessing in terms of a set of evaluation criteria and providing an overall normative argument about the relative importance of each of these criteria.

This leads to the formulation of two separate research questions for this article:

1. *How can we develop an argumentative framework for methodologically sound evaluative judgements in legal research?*
2. *How can we develop an argumentative framework for methodologically sound normative recommendations in legal research?*

3. State of the art: terra incognita

Over the past few years, there has been an intense debate on the methods of legal research.²⁰ This debate has been especially heated in the Netherlands and the Flemish region of Belgium.²¹ An important reason for this increased methodological awareness is that legal research has become more international, comparative, and interdisciplinary in character, and this requires that one's own methods be made explicit to allow for a productive dialogue across disciplines or jurisdictions.²² Moreover, the emergence of research schools and the organisation of methodology courses for LL.M. and Ph.D. students have led to various texts directed primarily to student audiences.²³ Methodological issues have probably been most widely discussed in the contexts of comparative law and empirical legal studies.²⁴

20 Hans-W. Micklitz, Guest Editorial. On the Politics of Legal Methodology, *Maastricht Journal of European and Comparative Law* 21 (2014) 4, 589: "In the last couple of years the market for academic legal books and law journals has been flooded with contributions on legal methodology."

For a general overview, especially of the US and European debate, see Van Gestel & Micklitz (note 5). For the debate in Australia, Canada, and the UK, see especially various publications by Terry Hutchinson, e.g. (note 3 and 16); Terry Hutchinson, *Researching and Writing in Law*, 2010; Watkins and Burton (note 16). For European perspectives, see Van Hoecke (note 16); Van Gestel, Micklitz & Maduro (note 8).

21 Most of these publications are in Dutch. A selection of relevant publications in English: Mark van Hoecke (2004) (ed.), *Epistemology and Methodology of Comparative Law*, 2004; Van Hoecke (note 16); Van Klink & Taekema (note 3); Jan M. Smits, *The Mind and Method of the Legal Academic*, 2012; Stolker (note 3); Sanne Taekema, Bart van Klink & Wouter de Been (eds.), *Facts and Norms in Law. Interdisciplinary Reflections on Legal Method*, 2016; Leeuw & Schmeets (note 14). There have also been various special issues of academic journals (e.g., *Erasmus Law Review* 2011, issue 1; 2015, issues 2 and 3; *Maastricht Journal of European and Comparative Law* 21 (2014) 4). A special bilingual journal *Recht en Methode / Law and Method* started in 2011; currently, almost all contributions are written in the English language.

22 Cf. Van Gestel, Micklitz & Maduro (note 8), 5.

23 Examples are Michael Salter & Julie Mason, *Writing Law Dissertations. An Introduction and Guide to the Conduct of Legal Research*, 2007; Mike McConville & Wing Hong Chui (eds.), *Research Methods for Law*, 2007; Terry Hutchinson, *Researching and Writing in Law*, 2010; Cryer, Robert Cryer, Tamara Hervey, Bal Sokhi-Bulley with Alexandra Bohm, *Research Methodologies in EU and International Law*, 2013; Watkins and Burton (note 16).

24 For comparative law, see Van Hoecke (notes 16 and 21); Mathias Siems, *Comparative Law*, 2014; Geoffrey Samuel, *An Introduction to Comparative Law Theory and Method*, 2014; Maurice Adams & Dirk Heirbaut,

Surprisingly, questions regarding evaluations and recommendations have so far largely been neglected in this debate. Only a few authors have explicitly taken up this theme.²⁵ Singer identifies the problem in a general way that includes academic research and teaching, but when he discusses methods these are restricted to normative arguments in the practice of lawyers and judges.²⁶ Smits argues that “the ultimate question of legal science is what the law *ought* to be”.²⁷ However, he remains vague as to how this could be done. He provides only part of the answer by arguing that comparative law may be used to find possible solutions and arguments²⁸ – in other words, it functions only as a mere heuristic – but fails to answer how these arguments may be weighed and evaluated.²⁹ The most explicit discussion is provided by Van Klink and Poort, who emphasise the normative character of legal research, and advocate value-based research.³⁰ However, apart from making a distinction between values and standards that are internal and external to law, and advocating (with only a summary argument) that legal scholarship in the narrow sense should only appeal to the former, the authors provide little guidance on precisely what value-based research should entail.

Of course, legal researchers frequently provide evaluations and recommendations. However, these studies focus only on a few values and other standards that the law is supposed to realise, and ignore most other values that might be relevant.³¹ Empirical legal studies show what the effects of a law are, and may sometimes provide plausible predictions about the effects of proposed statutory changes.³² Normative law and economics provides insights into the costs and the impact on wealth maximisation. Doctrinal research can analyse whether the law is coherent and in line with fundamental principles of the legal order, while philosophy may evaluate law in the light

The Method and Culture of Comparative Law. Essays in Honour of Mark van Hoecke, 2014. A strong version of the argument that legal researchers should follow the rules of empirical research whenever they make empirical claims may be found in Epstein & King (note 6). For a criticism that this model does not work for primarily doctrinal and normative research, see Jack Goldsmith & Adrian Vermeule, *Empirical Methodology and Legal Scholarship*, *University of Chicago Law Review* 69:1 (2002), 153–167.

25 Apart from the three publications discussed in the text, see esp. Leeuw & Schmeets (note 14), 220–235; Salter & Mason (note 23), 100–108.

26 See Singer (note 2), 905: “Both law professors and students are in need of advice about how to think about the nature of morality, fairness, and justice. More importantly, they need vocabulary for talking about normative matters and a set of resources and methodologies for structuring relevant arguments.”

27 Smits (note 21), 41

28 This idea clearly resembles the suggestion by the US Supreme Court Justice Louis Brandeis in *New State Ice Co. v. Liebmann* 285 U.S. 262 (1932) that states in a federal system can function as laboratories of democracy.

29 For a similar critique, see Vranken (note 4), 153–4 and 195–6.

30 B. M. J. van Klink & L. M. Poort, *De normativiteit van de rechtswetenschap. Een pleidooi voor meer reflectie op de normatieve basis van het recht en de rechtswetenschap*, *RM Themis* 6 (2013), 260

31 Some authors claim to provide an all-things-considered evaluation by translating and reducing certain values to one common value, or simply by declaring other values irrelevant. An example is the attempt by some law and economics scholars to focus only on welfare maximisation, and trying to put prices on non-monetary goods. For critiques on this reductionist approach, see Jules Coleman, *The Grounds of Welfare: Fairness Versus Welfare*. By Louis Kaplow & Steven Shavell, *Yale Law Journal* 112 (2003), 1511–1543; Renny Reyes, Alessandro Romano & Cecilia Emma Sottolotta (2015), *Regulatory Impact Assessment in Mexico: A Story of Interest Groups Pressure*, *Law and Development Review* 8 (2015) 1, 99–121.

32 For *ex ante* evaluative studies of legislation, see Verschuuren (note 16).

of normative theories. Each of these approaches provides relevant building blocks for a complete evaluation and for recommendations, but not a full argument. These studies are important in their own right, but insufficient as the basis for a complete, all-things-considered evaluation.³³

The limitations are illustrated by empirical evaluation studies, which usually take as the standard for evaluation the purposes of a statute (or more broadly a policy programme), and investigate whether these have been effectively realised.³⁴ Most studies evaluating legislation, executed or at least initiated by government agencies, have a similarly narrow focus on realisation of the goals of the statute, and include only a limited range of possible other effects.³⁵ In the Netherlands, for example, the emergence of legislative science has led to the formulation of “general principles of proper legislation” and to various suggestions for lists of criteria on which to evaluate legislation.³⁶ However, these criteria focus primarily on technical-legal and instrumental aspects, and do not address the substantive political dimension of legislation. The same is true for Regulatory Impact Assessments in the context of the EU associated with the Better Regulation project.³⁷ As a result, evaluation studies may identify certain *pro tanto* deficiencies in the law on the basis of certain relevant values, but do not make the leap from those partial deficiencies to a full all-things-considered evaluation, let alone to all-things-considered normative recommendations.

To conclude, we may find relevant building blocks and partial insights in legislative evaluation studies and other disciplines. However, there are two major deficiencies in these studies. The first is that they focus on only one or a restricted set of standards of evaluation. The second is that they do not provide methods for integrating into a broader interdisciplinary perspective the partial insights produced by various disciplines. Thus there is no integrated framework that can justify all-things-considered evaluations

33 For similar critiques on the limitations of various types of legal research that try to avoid fully addressing normative issues, such as doctrinal research and economic analysis of law, see Singer (note 2), 913 ff.

34 Two examples: In the widely used handbook Michael Quinn Patton, *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 2015, 18 the question of programme evaluation (of which legislative evaluation is a subtype) is formulated as “whether the people involved are accomplishing what they want to accomplish.” Only a few other minor remarks are made about the underlying values throughout the 806-page book; there are not even index entries for values or norms. In his case study on a ban on smoking in cars carrying children, Pawson (note 19), 174–5 only focuses on studying empirical aspects of the ban. Evaluation should, in his view, be based only on evidence, but he pays no explicit attention to how to go from facts to recommendations.

35 In the Netherlands, many new statutes contain explicit requirements to evaluate them after a number of years. Veerman (note 19), 13 estimates that 10 % of the yearly production of statutes in the Netherlands contains an explicit clause requiring evaluation. In Veerman’s meta-evaluation of 67 of these evaluation studies, the research question it concerns is simple: namely, whether the statutes have realised their goals. Side effects are only indirectly included insofar as they are relevant to the realisation of these goals.

36 See e.g. P.J.P.M. van Lochem (2015), *Kwaliteit van wetgeving als keuze*, in: *Kwaliteit als keuze – kwaliteit (sbeoordeling) van rechtspraak, wetgeving en rechtswetenschappelijk onderzoek* (Handelingen NJV), 2015, 141–242, esp. 150–167; Veerman (note 19); Willem J. Witteveen, *Alternatieve regulering. De vele gezichten van de wetgever*, in: *Alternatieve regelgeving* (Handelingen NJV), 2007, 1–66; Willem Witteveen, *De wet als kunstwerk. Een andere filosofie van het recht*, 2014.

37 For the EU see *Principles of good regulation*, 2003; Anne C. M. Meuwese, *Impact Assessment in EU Lawmaking*, 2008; Verschuuren (note 16), 3.

and recommendations. For this more holistic aim, the academic literature provides no guidance.³⁸

Sceptics may argue that such an integrated framework is impossible and undesirable, and that empirical or doctrinal legal researchers should simply refrain from providing evaluations or recommendations.³⁹ It is not seen as their job, because there is no adequate methodology that would allow them to execute it properly. Such an approach, in line with traditional positivist theories of science, is a respectable position.⁴⁰ However, there are good reasons not to accept it. First, the conclusion that no sound methodology exists is too hasty, as long as we have not even seriously tried to construct one. Second, legal researchers have been making evaluations and recommendations for a long time, based partly on practical experience and implicit knowledge, and partly on theoretical insights and research. Even if the methods can be improved, there is no reason to assume that the current evaluations and recommendations are so bad that they are worthless. Third, if legal scholars were to refrain from making evaluative judgements and recommendations, to whom then should society turn for expert advice on legal reform? If legislation is not merely a matter of whims and political preferences, we should expect academic researchers with some expertise in the specific field of law to be able to contribute to a critical analysis of legislation and to explore the possibilities of legal change.⁴¹ It seems unjustified to exclude from discussions on legal reform those with expert knowledge of the subject. It would be like saying that because medical doctors make mistakes, we should leave medicine to lay people altogether.

Even so, there is an important warning inherent in the sceptical position, and it is that not all legal research should be oriented towards evaluations and recommendations. Indeed, I will argue below that researchers should refrain more often from making evaluations and giving recommendations, not only because of an inadequate methodological basis but also because it is not always appropriate to do so. Purely expository doctrinal

38 It is remarkable that in Ph.D. and LL.M. student guides for doing legal research these broader evaluative and normative dimensions are at best discussed in only a few scattered remarks. See e.g. Hutchinson (note 23), 176; Cryer, Hervey, Sokhi-Bulley & Bohm (note 23), 10. Some guides even don't mention them at all, and restrict evaluation to empirically determining whether a law is effective; examples are Watkins and Burton (note 16); McConville & Chui (note 23). The most extensive discussion is provided by Salter & Mason (note 23), 100–108.

39 For a defence of such a position, see e.g. Anne Ruth Mackor, *Legal doctrine is a non-normative discipline: An argument from abstract object theory*, in: Taekema, Van Klink & De Been (note 21), 127–149. A relativist version of this argument is that, as values – and consequently evaluations – are mere preferences, it is wrong for lawyers or researchers to impose their views on others. Of course, as Singer admits, “we have no iron-clad way to prove a claim of morality or justice.” Joseph William Singer, *Critical Normativity*, *Law and Critique* 20 (2009), 27–42, 31. However, that sometimes only a weak justification is possible for normative claims does not imply the relativist conclusion that every attempt to justify is arbitrary.

40 It is respectable, but as a pragmatist I reject the presuppositions of this positivist view of science. Elsewhere I have defended the claim that an integration of doctrinal, empirical, and normative analyses is possible and desirable. (See Taekema & Van der Burg (note 12).) However, because I want to present a neutral framework here, I will leave aside this discussion on the merits of positivism.

41 In Wibren van der Burg, *The Need for Audacious Fully Armed Legal Scholars: Concluding Reflections*, in: Taekema, Van Klink & De Been (note 21), 265–286, I have argued that legal scholars should more often take up their responsibility to give advice involving their field of expertise.

work and purely empirical socio-legal research are important in their own right, and need not be followed at all by evaluative and normative judgments. Nevertheless, I also want to defend the position that sometimes it is possible and desirable to make justified evaluations and recommendations, and that, if we want to do so, we should try to do it in as methodologically sound a manner as possible.

4. The Characteristics of Evaluative and Normative Arguments

Legal research is an argumentative practice. Researchers put forward arguments to support their claims, and present counter-arguments to undermine the claims of other researchers. For example, they refer to legislation and case law but also to normative principles to defend their reconstruction of doctrine. They use experiments and observations as well as references to generally accepted theories to support their empirical claims. This argumentative character is very explicit when it comes to evaluations and recommendations. Such arguments are not a matter of mathematical proof, providing a strictly deductive line of argument.⁴² However, they are also not empirical proof, as they cannot rely – at least not fully – on relatively strict methods for establishing facts and theories based on facts. I suggest that at least the following characteristics of arguments in the context of legal research should be taken into account.⁴³ Most of these characteristics hold for arguments in academic research in general, but some of them are more specific to arguments that lead to evaluations and normative recommendations.

A. Arguments are addressed to an academic forum

The community of scholars is the primary forum for assessing the research,⁴⁴ and expert peers should decide on the quality of the research and of the arguments. The only – albeit imperfect – way to judge whether arguments in legal research are sound and convincing is to submit them to this type of forum, and to see how well they can stand up to

42 Cf. Singer (note 39), 28.

43 I focus on such arguments in the context of legal research, although much of what is said here could also apply to arguments in the context of legal practice as well as in various types of non-legal disciplines such as ethics. There is an immense body of literature on legal reasoning in general to which I cannot do justice in this context; obviously, much of it would be relevant for a more detailed analysis of arguments in the context of legal research.

44 Cf. Vranken (note 4), 37 ff. Various groups of legal practitioners – judges, lawyers, legislators – may also constitute relevant fora for the presentation of legal research. They may assess the research on whether it is convincing and useful for their practical purposes, but this is not a criterion of quality or soundness of the research as academic research. For example, in this assessment by members of the legal practice, there may be a bias against research that fundamentally criticises the legal order – and therefore may be regarded as of little practical use – and towards research that suggests practical solutions and quick fixes for concrete problems. Moreover, these ‘consumers’ of legal research are usually not much interested in whether these practical suggestions are based on methodologically sound research. The question of whether the research is sound should be assessed by the academic forum.

that trial. The method is imperfect because the forum of course can be wrong,⁴⁵ but even then, the researcher can claim that an ideal forum would accept his arguments.

To enable the forum to assess the arguments, the researcher cannot simply report the results of her research. She has to spell out each of the links in the chain of argument. Of course, this is a basic requirement for every type of research, including doctrinal research. Did the researcher arbitrarily exclude data (such as specific cases or statutes) that do not support her conclusion? Did she make clear why in hard cases she chose one solution rather than the other? This is even more important in evaluative and normative arguments. The forum must have all arguments and counter-arguments on the table. Basically, this role of the forum is the reason for various other characteristics and requirements.

B. Criteria for assessing arguments are soundness and convincingness

We should make a distinction between *sound* and *convincing* arguments. Disagreement is inherent in law and in legal research – as in other disciplines such as philosophy.⁴⁶ Even the ideal forum need not fully accept the argument as convincing.⁴⁷ The primary standard for determining the quality of arguments is not weaker but it is different: the argument must be sound, that is, it must be logically valid and it must be defensible.⁴⁸ It is like many interesting cases argued in supreme courts: even if one side loses in the end, it is worth trying because there is a defensible case. Moreover, the court may turn out to be wrong in the long run. When editors or supervisors have to decide whether to accept an article or a dissertation, the criterion should be soundness.⁴⁹

Of course, when a reader of a journal article decides to write a critical response, she will focus both on soundness and convincingness. A researcher hopes that the academic forum will accept the argument as both sound and convincing. However, convincingness has a more subjective dimension that may vary with the audience. For example, a supervisor may accept a Ph. D. candidate's critique on her own views as a defensible and highly original one, and therefore accept the dissertation as sound research fully deserving

45 To avoid misunderstanding, I should mention that the forum's acceptance of an argument as sound is not the *criterion* or the ground for its soundness. A researcher with a brilliant mind ahead of its time may be unable to convince the forum that her argument is sound. Nevertheless, it is part of the practice of scholarship and research that we present our research to the forum in such a way that it can be critically assessed.

46 Jeremy Waldron, *Law and Disagreement*, 1999; Raymond A. Belliotti, *Justifying Law. The Debate over Foundations, Goals, and Methods*, 1994

47 If only because the academic forum includes a diversity in disciplinary, theoretical and methodological backgrounds.

48 Eveline Feteris & Harm Kloosterhuis, *Law and Argumentation Theory. Theoretical Approaches to Legal Justification*, in: Van Klink & Taekema (note 3), 255 define a sound argument as being a logically valid one with acceptable premises. Instead of acceptable, however, I prefer the weaker notion of defensible.

49 Both convincing and sound are vague concepts that may give rise to controversy. Whether an argument is considered sound or convincing, depends partly on the context and on the audience. However, that there is a grey zone of vagueness and controversy does not imply that we can do without the distinction between sound and convincing.

ing of a doctorate, but still reject the critique as unconvincing. There are many academic debates in which both parties fully understand each other's positions, even regard them as respectable and defensible, but are not convinced by the arguments.

C. Arguments are embedded in controversies

Law is characterised by differences in opinion. This holds for legal practice but also for legal research. Of course, on many issues there is – or seems to be – a broad consensus, but research often focuses on points of controversy or on points where the matter is still open and undecided. The controversial character of arguments should be fully acknowledged. Therefore, an academic publication should not merely put forward views supporting the conclusion. It should also mention alternative views, represent them impartially according to the principle of charitable interpretation, and then make an argument as to why these alternative views are not, or not fully, convincing. Only in this way can members of the academic forum fully assess the argument.⁵⁰

D. Arguments are defeasible⁵¹

Evaluative and normative arguments rarely lead to knock-down conclusions. Usually they make a conclusion merely plausible: for example, the suggestion that a tax rule should be amended.⁵² The conclusions are not based on strict deductive reasoning, but on defeasible reasoning. The conclusion is plausible if the reasons pro outweigh the reasons contra. Even so, these conclusions are only provisional; they can always be outweighed or invalidated by new arguments. In order for the forum to assess whether the argument should be accepted, all relevant reasons pro *and* contra must be discussed in the article, and it must be argued why these reasons are – or are not – sound and plausible. Moreover, it must be explained why the reasons against this conclusion are considered weaker than the reasons in favour.

50 See also Belliotti (note 46), 237. He argues that the fact of controversy does not exclude there being right answers in some cases; moreover, it is not a question of anything goes. He states: “Not just any world vision passes the test for a legitimate legal ideology, nor are we committed to the simple-minded proposition that one vision is no better or worse than any other.”

51 On defeasible reasoning, see Giovanni Sartor, *Defeasibility in Legal Reasoning*, in: Jordi Ferrer Beltrán & Giovanni Battista Ratti (eds.), *The Logic of Legal Requirements. Essays on Defeasibility*, 2012, 108–136

52 I use plausible here in a broad sense. It includes the degree of certainty for factual statements and the degree of acceptability for normative and evaluative statements.

E. Arguments are coherentist

Evaluative and normative reasoning is coherentist.⁵³ We do not merely apply a normative theory to concrete issues – the theory is also developed and refined in the light of moral intuitions and the relevant facts. In moral philosophy, John Rawls elaborated this basic idea into a method that he called reflective equilibrium.⁵⁴ A version of reflective equilibrium theory can provide an adequate description of what doctrinal scholars do when they construct legal doctrine.⁵⁵ More importantly for this article, I submit – but without elaborating further here – that reflective equilibrium can also model evaluative and normative reasoning in the context of legal research. This going back and forth takes place throughout the research process; right up to the end, theoretical conclusions may be refined in the light of concrete intuitions.

The academic publication, however, only presents the end result: namely, a coherent argument. This argument will often be presented in a one-directional way, in which references to facts, strong intuitions, normative principles and so on are all put forward as reasons to support the argument or to undermine and weaken counter-arguments. The dialectical process of discovery is ignored in the presentation. As a consequence, there may be significant differences between the context of discovery and the context of justification. In this article, I will focus on the latter: namely, on how results of the research process are presented to the academic forum. The distinct question as to how the process of discovery can be structured methodologically is left for future research.

F. Arguments require evaluation standards

In order to evaluate, we need evaluation standards. In legal research, these standards are often implicit and taken for granted. Common implicit standards for doctrinal scholarship are doctrinal consistency and coherence, and the fundamental principles or values

53 Of course, in this general formulation, this is a highly controversial statement, as there are many foundationalist normative theories that could be used in evaluating the law. I have argued for coherentist justification in various publications (e. g. Wibren van der Burg, *Dynamic Ethics*, *Journal of Value Inquiry* 37 (2003), 12–34), but cannot repeat the analysis in this article. Foundationalist ethical theories may be used for an external criticism of law, but usually only lead to a *pro tanto* evaluation, based for example on the value of efficiency or justice. In order to combine all relevant *pro tanto* evaluations into an all-things-considered evaluation, we need to balance and integrate them, and this can best be done in a coherentist process of reasoning. Most importantly, in order to produce sound recommendations, evaluations have to be translated and incorporated in an internal legal perspective. The internal perspective of law in which a doctrine is critically reconstructed and suggestions are made to change the law requires a hermeneutic approach, which implies coherentism.

54 John Rawls, *A Theory of Justice*, 1971. On reflective equilibrium, see Norman Daniels, *Justice and Justification: Reflective Equilibrium in Theory and Practice*, 1996; Wibren van der Burg & Theo van Willigenburg, *Reflective Equilibrium*, 1998.

55 For reflective equilibrium as applied to law, see Ronald Dworkin, *Taking Rights Seriously*, 1978, 160 f.; Marsha P. Hanen, *Justification as Coherence*, in: Michael Alexander Stewart (ed.), *Law, Morality and Rights*, 1983, 67–92; Singer (note 2), 976.

of the legal order. In law and economics, the standards are usually variations of welfare, preference satisfaction, and efficiency. For socio-legal studies, the standards may be the purposes of a statute as well as the derived value of effectiveness in realising those purposes. However, apart from these familiar standards, there may be many more that are relevant to evaluation. The choice of standards raises many difficult questions, which I will discuss in Section 5.

G. Arguments are complex

Arguments in legal research are usually complex, consisting of a set of reasons and auxiliary reasons for and against reasons. The links in a chain of argument can be quite complex and long: for instance, “This statute is a good law because A, B, and C; A is plausible because X and Y and despite Z, and so on.” There must be a full chain of argument, in which each of the links is made explicit and then made plausible. The argument is as strong as its weakest link. A famous cartoon by Sidney Harris shows a complex mathematical proof with formulas on a blackboard, in which one of the links is “then a miracle occurs”. The comment of a colleague is “I think you should be more explicit here in step two.”⁵⁶ This cartoon provides a cautionary tale for evaluative and normative arguments. No step should be taken for granted; every step should be made explicit and made plausible. If there are missing or implausible links, the whole argument may fall apart.

H. An argument for normative recommendations is comparative

Arguments for normative recommendations are comparative in two respects. First, they must compare the suggested solutions to the status quo. It should be demonstrated not only that the alternative is better according to one evaluation standard but better all-things-considered. Legal reforms are no zero-sum games, but an improvement in one evaluation standard may sometimes lead to deterioration in terms of another standard. For example, elaborating a statute in more precise and legally correct words may result in a higher degree of consistency, but could decrease compliance because this technical language is so complex that ordinary citizens no longer understand the rules. Second, they must compare the suggested solutions to possible alternative solutions. Again, this should be an all-things-considered assessment, because it is likely that different solutions will score higher on different dimensions.

56 <http://www.sciencecartoonsplus.com/pages/gallery.php> (last visited on 22 February 2018)

I. Arguments may appeal to authorities or generally accepted ideas

John Rawls has argued that “justification proceeds from what all parties to the discussion hold in common.”⁵⁷ In principle, every reason given in an argument can be met with a further “Why?”⁵⁸ There is usually a point where truths are held to be, if not self-evident, at least so generally accepted that further argument may be deemed superfluous: for example, “Torture is wrong”, “Free speech is a fundamental human right”. For practical purposes, such appeals to generally accepted ideas or – as Rawls calls them – provisionally fixed points will usually suffice.⁵⁹ Lawyers use a specific type of provisionally fixed points: namely, legal sources. Examples of these are references to a recent Supreme Court decision or to a recent statute. In the world of Civil Law, an authoritative handbook may also provide a provisionally fixed point. Lawyers are accustomed to such appeals to authority, whereas this type of argument outside legal contexts is often regarded as fallacious or at least suspect.

However, even in expository legal research such an appeal to authoritative sources must be treated with caution.⁶⁰ After all, a supreme court can overrule its precedent, an author might be mistaken, or an unforeseen interpretation of a statute may emerge. For evaluative and normative arguments, the appeal to authorities is even more questionable. John Rawls may be the most influential political philosopher of the last century, but this is not sufficient reason to accept appeals to his theory.⁶¹ If we apply his theory, we must provide further reasons why it is adequate in this specific context. A similar point can be made with an appeal to judicial interpretations of specific concepts or arguments. For example, the fact that the US Supreme Court (in *Bowers v Hardwick*) excluded homosexual relations from protection by the right to privacy was even in 1986 not sufficient reason for a legal scholar to accept the majority’s conception of privacy.⁶² After all, plausible alternative conceptions had been suggested by the Court’s minority, in the literature and by other courts: for instance, by the ECHR in the *Dudgeon* case.⁶³ If

57 Rawls (note 54), 580

58 See Singer (note 2), 903: “We know the feeling of being confronted with the ‘why?’ question and not knowing what to say after ‘because ...’, yet feeling we should be able to say something.”

59 Rawls (note 54), 20. See also Singer (note 2), 966: “Evaluative assertions can provide an answer to the ‘because clause’ by appealing to the things we in fact already believe.”

60 See also Taekema (note 3), 48.

61 Brian Leiter, Intellectual Voyeurism in Legal Scholarship, *Yale Journal of Law & the Humanities* 4 (1992), 79–104, 92 rightly criticises “the startling absence of *argumentation* in most discussions of philosophical issues” in law journals. Too frequently, authors merely quote statements by philosophers rather than provide arguments for why these statements are correct (or why the author believes it is justified to restrict the analysis to these philosophers). In my view, it is certainly not unique to legal scholarship to put “citation in place of argument” (Leiter (this footnote), 92); it is quite common among trained philosophers as well.

62 *Bowers v. Hardwick*, 478 U.S. 186 (1986)

63 *Dudgeon v. United Kingdom*, 22 October 1981, No. 7275/76. Of course, the fact that this interpretation of privacy was later rejected in *Lawrence v. Texas*, 539 U.S. 558 (2003) illustrates clearly that academic researchers should be very cautious with regard to grounding the standards for evaluative analysis in the authority of court decisions.

we want to evaluate court cases or legislation critically, judgements and statutes may be sources for arguments and interpretations, but they are not authorities.

5. Standards of Evaluation

As mentioned in the previous section, legal researchers use a variety of standards for evaluative purposes. Take the example of a 10 % tax increase on cigarettes. On the one hand, there are very specific evaluation criteria, such as changes in tax revenues, in the numbers of smokers, and in the average consumption of cigarettes per smoker. Although it may not always be easy to gather the data, these evaluation criteria are clear and specific. On the other hand, there are also quite general values, such as public health and individual freedom. Although they are obviously relevant to the evaluation of the tax increase, it is not easy to interpret and operationalise them, let alone to determine the implications of the tax raise for these standards. Both for the specific and general standards, a crucial question has to do with how much importance they should have in an all-things-considered evaluation.

This example illustrates various problems. First, some standards are directly valuable, such as public health and freedom, whereas others are indirectly valuable, such as increased tax revenues or a decreased number of smokers. We need additional arguments as to why they are important. Second, evaluation standards are diverse in terms of specificity and generality. Third, it may be difficult to interpret and operationalise the more general standards and to determine how important they are. For example, what does freedom or public health mean? Is public health important as such and not merely in terms of health care costs? Fourth, if we have to choose because we cannot research everything, which standards are the most important? Fifth, how can we combine this diversity of standards in an all-things-considered evaluation or recommendation?

My suggestion is that these questions can best be addressed if we assume that standards of evaluation refer directly or indirectly to fundamental values. As the word evaluation suggests, we need to assess whether and to what extent certain values have been realised in order to determine whether the law is good. Thus, the notion of values can provide a common conceptual framework to integrate the various evaluation standards. However, a focus on values as the common denominator is only a first step in addressing these five issues.

First, we should make a distinction between direct and indirect standards. Some evaluation standards are based only indirectly on values. For example, doctrinal consistency is clearly a standard for evaluation, but it is at most an instrumental, dependent value rather than an ultimate value. We need further arguments as to why consistency is valuable, which can be found in values such as the rule of law, equal treatment, or legal certainty. The same holds for effectiveness. Effectiveness is only an instrumental value insofar as the goals served by the law are valuable. We would not value effectiveness if a law had immoral purposes – for example, the suppression of free speech – we would prefer instead that the law be quite ineffective. Many evaluation standards may need a

process of transformation in order to be understood in terms of underlying values. How valuable is it that tax revenues increase and that numbers of smokers decrease?⁶⁴ Again, for a full evaluation, researchers should make explicit why certain evaluation standards are important, and especially how significant they are compared to other standards. Especially for this latter aspect, the issue of balancing, it is important to have a common denominator of values.

Second, we need to address the question of how to translate or relate the variety of standards to values. I suggest that – as subcategories of the broad category of evaluation standards – a distinction between values, conceptions of values, and criteria may be helpful here. Values such as freedom or consistency can be interpreted in various, conflicting ways; in the terminology of Rawls and Dworkin, these concepts are open to a variety of conceptions.⁶⁵ These more specific conceptions usually cannot be applied directly to concrete cases, but must be specified or operationalised as criteria in order to be useful for testing and evaluating. A researcher must explicate the conception she chooses and justify the choice, as well as explicate and justify the choice for specific criteria.

Although the threefold distinction I present here moves from the more abstract to the specific, that is not always the order of analysis in real research. On the one hand, empirical researchers frequently start with concrete criteria (such as tax revenues and numbers of smokers) because that is what they can find with the help of their empirical methods. However, if they want to use their results for all-things-considered evaluations and recommendations, they must analyse and translate these in terms of underlying values, and explain which conceptions of these values they use and why. On the other hand, philosophers usually focus on more abstract values such as justice or freedom, and have to go from these to conceptions and then to operationalisations.

This brings us to the third issue: namely, that of interpretation and operationalisation. Values can give rise to different interpretations or conceptions. For example, the value of legality can be interpreted in terms of the eight principles of Lon Fuller, or of the eight principles of Joseph Raz, but it can also be interpreted to refer to procedural values and substantive values.⁶⁶ This variation is not something that will simply disappear after elaborate conceptual analysis, as values are essentially contested concepts.⁶⁷ This means that they are inherently open to a variety of partly overlapping, partly incompatible conceptions.

64 To add even more complexity: these questions can not be answered in general, but vary with the context. For example, how important the increase in tax revenue is, depends on the budget deficit.

65 Rawls (note 54), 5; Dworkin (note 55), 103 and 134

66 Lon L. Fuller, *The Morality of Law*, 1969; Joseph Raz, *The Authority of Law. Essays on Law and Morality*, 1979, 210–229; Jeremy Waldron, *The Rule of Law as a Theater of Debate*, in: Justine Burley (ed.), *Dworkin and his Critics*, 2004, 319–336. For an overview see Jørgen Møller and Svend-Erik Skaaning, *The Rule of Law. Definitions, Measures, Patterns and Causes*, 2014

67 See Wibren van der Burg, *Law as a Second-Order Essentially Contested Concept*, *Jurisprudence* 8 (2017), 230–256, in which I argue that the value orientation of essentially contested concepts is crucial in explaining the essential contestedness.

For instance, there are many different interpretations of democracy, ranging from direct democracy to proportional representation and a district system. If we want to examine whether a bill introducing a referendum improves the democratic quality of our constitutional order, we must discuss the various conceptions of democracy, and give arguments as to why we choose one specific interpretation. As the question is one of improving our existing political system, we must start with those theoretical conceptions that fit our order, or that are at least not too much in tension with it. This will restrict the range of plausible conceptions, but there still may be various conceptions that meet that criterion. For a full evaluation, it may be necessary to assess the referendum in light of each of these conceptions.

This example illustrates that we should sometimes choose more than one conception, and determine whether the bill is consistent with each of them. We may call this conceptual triangulation. Whereas methodological triangulation uses different methods to gain a richer insight into a phenomenon, conceptual triangulation uses different conceptions to obtain a more comprehensive overview as to whether a phenomenon is consistent with a specific value. In most cases, however, we will have to choose one specific conception to make assessment feasible. Again, although there are no easy guidelines here, we should argue why we have made this choice.

A connected problem is that of operationalisation.⁶⁸ Once we have chosen a specific conception, we must determine how to assess it by formulating more specific criteria. It may not be measurable in a quantitative way, but we must be able to judge objectively whether or not the law fails to meet the standard.⁶⁹ Of course, the availability of methods may restrict which criteria we are able to study. When we design a research project, there is usually a dialectical interplay between available methods and possible criteria. We determine in the light of possible methods what criteria can be feasible and vice versa. However, when we present the results, we should begin with the more general values and conceptions, and then explain the specific criteria and the methods that have been used to study them.

Operationalisation does not merely require a more specific standard; it also requires that we decide what degree of realisation would be desirable and realistic.⁷⁰ Because we live in an imperfect world, a full realisation of values is not possible. Nor, as we may learn from Radbruch and Fuller, is it always desirable, because the full realisation of one value may lead to perverse effects and to a lesser realisation of other important values.⁷¹ Values can conflict. It would not be fair criticism of a law against murder that murders

68 The notion of operationalisation is well known in social sciences, but I suggest that it is useful as well in the context of doctrinal research. The construction of a more elaborate conception of a value is basically a theoretical exercise, whereas deciding how to assess or measure it for specific purposes is basically a methodological issue; it partly depends on the available methods as to how we should operationalise the value.

69 For a discussion of how to elaborate the concept of the rule of law in conceptions and operationalisations, see Møller & Skaaning (note 66).

70 This need not always be done in advance, but it should at least be discussed explicitly when the results are presented in a publication.

71 For the unavoidability of conflicts or antinomies between the core ideals of law but also between different aspects of one ideal, see Fuller (note 66), 41 and 45; Gustav Radbruch (1950 [orig.1932], *Legal Philosophy*,

are still taking place and therefore the law does not work. The same holds for laws combating, for example, discrimination or environmental pollution. Perhaps for a superficial evaluation it may seem enough to show that a certain value is not completely realised. However, we may ask what is the worth of such a limited evaluation. We know that we do not live in a perfect world, and that our laws are in many ways not completely coherent, just, democratic, or sustainable – they are often flawed in the light of whatever substantive values we would want to use for evaluation. This shows that we have either to set reasonably attainable cut-off points or to argue why it is not possible to do so, and merely designate a grey zone of somewhat acceptable degrees of realisation.

Fifth, there is a broad range of possibly relevant values. Due to practical limitations of time and disciplinary competence, most legal researchers use only a small set of general values or concrete criteria. For example, they focus on doctrinal consistency and coherence, combined with an analysis of whether a statute or judicial decision can be justified in the light of fundamental legal principles.⁷² This may suffice for a *pro tanto* evaluation, but even then the question arises as to how can researchers determine on which values and criteria they should focus, and how can they justify their choice. Nevertheless, for a full, all-things-considered evaluation, and especially for useful normative recommendations, we need to include all possibly relevant standards.⁷³

Obviously, this includes distinctly legal values, such as legality and legal certainty, but the range is much wider. In principle, every value can be relevant for the evaluation of law. They need not be broadly accepted – controversial values may also be used. For example, the researcher may include values such as marital fidelity or animal integrity if she can make a defensible case for them. After all, the test of a good argument is not whether it is convincing but whether the academic forum considers it sound.⁷⁴ Researchers did not have to wait until there was a broader acceptance of homosexuality before they could appeal to an inclusive conception of equality to criticise legal discrimination on the basis of sexual orientation. This example illustrates that the set of values – and the set of conceptions of those values – that may be used in evaluation is open, and that it may change. It seems likely that our views with regard to the environment and animals will change in the future. New values may emerge, and attempts to develop and use such emerging values should not be discarded because they are currently controversial.

This open character of the set of possibly relevant values presents major problems for researchers. There is no simple checklist that they can tick to see whether a law is

in: *The Legal Philosophies of Lask, Radbruch, and Dabin*, (ed. E. W. Paterson (1950), K. Wilk transl.), 1950, 109–112; Sanne Taekema, *The Concept of Ideals in Legal Theory*, 2003, 181–183.

72 On coherence, see Nicola Lacey, Contingency, Coherence, and Conceptualism: Reflections on the Encounter between “Critique” and “the Philosophy of the Criminal Law”, in: *Philosophy and the Criminal Law. Principle and Critique* (ed. Antony Duff), 1998, 9–59; Salter and Mason (note 23), 100; Amalia Amaya, *The Tapestry of Reason. An Inquiry into the Nature of Coherence and its Role in Legal Argument*, 2015.

73 A further issue is how to deal with the conflicts between those standards. Some philosophers, but certainly not all, hold that values are incommensurable; consequently, there is often no decisive argument to prefer one solution over another. As my aim in this article is to provide a neutral framework, I will leave the issue of incommensurability aside.

74 This does not mean that anything goes; these controversial values must still be minimally defensible.

justified in view of the underlying values. More importantly, if they want to restrict their analysis to only one or two values (for example, fairness and efficiency in the evaluation of tax law), they have to justify this choice by making plausible that other values are less relevant. Alternatively, they can stick to a mere *pro tanto* evaluation, but then they should mention this restriction explicitly as well. There is no easy solution to this problem; the only feasible guideline is that no choice, no restriction can be taken for granted but must be justified as thoroughly as possible.

6. The Chain of Justification: The Structure of Evaluative Arguments

A recurrent theme in this paper is that we should make every step in the argument explicit, and that we should justify each one. In order to identify the most important steps, I develop in this section a basic structure for presenting evaluative arguments. By submitting them in a linear fashion, I present them in the context of justification. In reality, legal research is often hermeneutical and circular, going back and forth between the different steps until a coherent story can be told.⁷⁵ The various steps are thus not linear, but are tested and refined during that process. The context of discovery may therefore be quite unsystematical, but when we want to justify evaluative conclusions to the forum, the arguments must be presented systematically. I suggest that we should present the justification as a clear chain of links in which each link is made both explicit and plausible. A full evaluative argument should at least address the following six questions.

A. What is the precise subject of evaluation?

Research projects usually focus on only a small part of the legal order: a court case, a statute, or a specific subfield of law. This is not a neutral choice. A biased evaluation results, for instance, if we gauge the fairness of one specific tax rule that provides deductions that mostly benefit the rich but we do not include the broader tax structure in which the rich pay progressively higher tariffs. The same would hold, of course, but with a different bias, if we were to assess only the progressive tax structure without acknowledging that the rich usually profit more from deductions than the poor. The choice of topic and its demarcation should therefore be discussed explicitly, and justified in the light of what consequences the specific demarcation might have, especially in terms of biases.

75 With regard to the hermeneutical character of legal research, see Lacey (note 72); Taekema (note 3), 45–48.

B. What are the most important values to be used in the evaluation of this specific subject?⁷⁶

Which values are relevant is partly relative to the field of study. In tax law, economic efficiency may be a very important instrumental value but it is less so in most areas of constitutional law. In both cases, however, we should discuss not only efficiency but also the underlying values for which efficiency is instrumental, such as economic growth – an instrumental value as well, in turn referring to intrinsic values such as individual happiness. We should not restrict the set of values arbitrarily. A small research project usually does not allow us to include all relevant values, and probably not even the most important ones. However, if we do make a choice we should justify it. If the choice can affect the soundness of the evaluative conclusion, this should be mentioned explicitly and its implications should be discussed.

C. Which conceptions of these values are chosen?

As discussed above, each value may lend itself to different conceptions. Researchers should discuss the most plausible of these, and argue why they have chosen to focus on a specific one, or – in the event they use conceptual triangulation – on those specific ones.

D. How can these conceptions be operationalised into evaluation criteria, and what methods will be used to assess them?

The criteria need not be quantifiable. The decision to opt for qualitative research can be, of course, perfectly legitimate; in doctrinal research, quantitative methods even play a minor role. Even so, the criteria must be specific enough to make an objective assessment possible. Again, the choice of these criteria and methods must be justified explicitly. Moreover, as suggested above, we need to make explicit the degree of realisation that would suffice to place the law above criticism – even if this is a grey zone rather than a clear cut-off point.

E. How should the subject be evaluated in view of each of the criteria?

This, of course, is the proof of the pudding, since after choosing the evaluation criteria and the methods, we need to determine whether these criteria have been met. This must be done according to the methodological standards of the disciplines used for the

⁷⁶ To avoid misunderstanding, I should repeat that going from values to criteria is simply the order of presentation. In research, the hermeneutic process is a circular interplay between choosing values and choosing criteria.

evaluation, and can therefore vary widely. Although this is one of the most important links in the argument – and in most articles will even comprise the bulk of the text – I will not discuss it further here. Because it depends on the methodologies of each of the disciplines involved, it would require a very detailed and broad discussion.

F. How do we balance conflicting *pro tanto* evaluations?

Even if we have restricted the set of values and evaluation criteria, it is very likely that our evaluation will provide a mixed image. With regard to certain criteria the evaluation is positive, with regard to others negative. In those cases, we must discuss whether the overall balance is positive or negative, and make explicit the arguments for an overall judgement. This is made even more difficult, on the one hand, by the problem that values are often difficult to compare or even incommensurable and, on the other, by the contextual variation in the interpretation and the importance of values.

The academic forum can only assess whether the conclusions about the merits of the law are sound and convincing if each of the previous steps has been made explicit and the choices made at each of them have been made plausible. These five steps will usually result in a *pro tanto* evaluation of the law – *pro tanto*, as it is based only on the values that were included in the evaluation. The inclusion of other values – or other conceptions and operationalisations of these values – might have led to a different evaluation.

At first glance, a restricted *pro tanto* evaluation might be all we need. As I will argue below, it may be a legitimate aim of academic publications to provide this. However, on methodological grounds it will not always be possible to restrict our evaluation. Even for a *pro tanto* evaluation, we must determine how much realisation of the chosen values, as specified in concrete criteria, is required. For example, a high degree of legal certainty may lead to a suboptimal realisation of justice, or to a dysfunctional and outdated legal order that cannot keep up with societal changes.⁷⁷ In order to determine the optimal realisation of a certain value, we have to take into account the desirable and possible degrees of realisation of other values as well. Therefore, even when we want just a *pro tanto* evaluation, we may need to study other values – and the criteria derived from them – apart from the ones we have chosen as our primary focus.

A full, all-things-considered evaluation should – at a minimum – include judgements about the most important instrumental and intrinsic values, and thus must include analyses of doctrinal coherence and conformity with fundamental principles of the legal order such as legality, intended effects, side-effects, and costs as well as of the most important political and moral values at stake. A *pro tanto* evaluation may be presented as a full evaluation only if we can make plausible that we have included the most important values and criteria relevant for this subject, or that the inclusion of other values and criteria might not make a substantive difference. However, to make this plausible requires

77 This is a core thesis in the work of Radbruch, who holds that none of the three elements of the *Rechtsidee* can be fully realised, nor should we aspire to do so. Radbruch (note 71), 109–112.

at least a superficial discussion and informed guesswork about the most important values and their realisation, which is still a substantive burden of argumentative proof.⁷⁸

7. From Evaluation to Normative Recommendations

An evaluation may lead to the identification of concrete problems. This raises the question of whether the law can and should be changed in order to solve these problems. Sometimes the solutions may seem obvious. If doctrinal analysis shows that a rule in a statute is unclear, we should reformulate it. If a recent tax reform has unexpected unfair effects for particular groups of citizens, we should make the rules more just.

However, recommendations may not always be that easy. If there is an inconsistency between two rules, we may need a further argument to decide which of the two rules needs to be adapted. More importantly, solving one problem often creates another, as the promotion of one value may have a negative impact on other values. If we want to confiscate more weapons and drugs, providing the police with additional surveillance powers such as stop-and-search policies may seem effective. However, when these policies especially target members of minority groups in a discriminatory manner, they may also infringe upon civil rights and damage trust in the police.

Formulating methodologically sound normative recommendations usually requires a broad interdisciplinary perspective, because mere *pro tanto* recommendations have limited value and may even be misleading if not presented carefully with all due restrictions. In order to make fully justified recommendations for legal reform, even if the primary method was that of traditional doctrinal research, the possible effects and the normative-philosophical justification also need to be addressed. Perhaps they are not highly relevant or important, but if that is the argument not to include them, it should be made plausible, and that requires at least a minimal discussion of those dimensions.

This means that, even more strongly so than with regard to evaluations, justifying recommendations usually requires cooperation between at least three disciplines: doctrinal, empirical, and normative-philosophical. Each of these has major contributions to make, since recommendations will not be fully justified if one of them is missing. This does not exclude partial *pro tanto* recommendations, but we should be extremely cautious in using them as the basis for actual legislative reform.⁷⁹

A chain of justification similar to that described in the previous section can be constructed for the argumentative links leading to normative recommendations. The start-

78 For informed guesswork, see Pawson (note 19), 81.

79 This points to an important limitation to the usefulness not only of purely doctrinal approaches but also of empirical approaches such as law and economics or evidence-based legal studies. They can provide important partial insights, but these need to be supplemented by a broader analysis, including normative-philosophical and doctrinal research, in order to warrant all-things-considered evaluations and recommendations. When these types of empirical legal studies claim to provide more than a restricted partial evaluation, these claims do not take sufficiently into account the restrictions of the research.

ing point of this chain is that the evaluation has resulted in the identification of a clearly formulated problem.

A. Is the problem important enough?

Not every problem is important enough to be addressed explicitly. Legal change comes with a cost. If it happens frequently, it might weaken legal certainty and trust in the legal order.⁸⁰ Even a simple legislative change costs substantive amounts of time and money. Therefore, in order to justify normative recommendations that legal norms should be changed, it should be demonstrated clearly that there really is a serious problem that needs to be solved immediately.

B. What are the possible solutions to the problem?

Rather than already going for a seemingly obvious solution, we should pause to consider other possible alternatives. This is especially important because some alternatives may require thinking out of the box.⁸¹ The best solution with regard to an inconsistency between two rules need not always be to change one of them. The inconsistency could be an indication of a more fundamental tension. For example, animal law is often not fully coherent, because it has developed organically from earlier times when the interests of animals were largely ignored to one in which we are still discussing how to take animal welfare more seriously. We may expect further inconsistencies when introducing new animal welfare legislation. Perhaps the best approach to removing these inconsistencies requires a fundamental review of the basic principles of animal law, and that fundamental revision may affect rules that were at first glance not related to the existing inconsistency. In other situations, the solution might be to abolish legislation. Instead of opening up civil marriage to same-sex couples, for instance, marriage could simply have been abolished – as many feminists and gay-rights activists have advocated. This would have had the additional advantage that it would also have ended the unequal treatment of citizens who are single or live in polyamorous or multiple-parent arrangements. A weaker version of this would have been to eliminate all legal rules that differentiate between citizens on the basis of marital status. Of course, not all suggestions from outside the box need to involve such radical changes. Nevertheless, it may be useful to search for possible solutions that are less obvious.

80 One of Fuller's principles of legality is that of constancy through time (Fuller (note 66), 79–81). Raz (note 66), 213–214, formulates the same principle as the requirement that laws should be relatively stable.

81 According to Jonathan Verschuuren and Rob van Gestel, *Ex Ante Evaluation of Legislation: An Introduction*, in: Verschuuren (note 16), 5, often only the option preferred by the authorities is considered in *ex ante* evaluations, and there is generally “a lack of imagination as it comes to alternative ways of regulation.”

One way of doing so might lie in comparative research. Jan M. Smits has argued that comparative law may help us to find solutions and arguments that have been discussed in other jurisdictions, and to examine how they have worked there.⁸² We could also undertake an interdisciplinary comparison with ethics, for example, to gain inspiration from the ways in which ethical theories deal with issues such as responsibility, care, or organ transplants.⁸³

C. What are the most attractive and defensible solutions?

There will sometimes be many alternative solutions, in which case we must make a shortlist of those that are most attractive and defensible. To do so, we should scan all possible alternatives, select those that seem most attractive and defensible, and then justify our choice. The default option is always to do nothing, because the transition costs, the uncertainty as to whether the expected improvements can be realised, and the possibility of unexpected negative side-effects always provide good *pro tanto* arguments for not changing the law. Hence, the inaction alternative should always remain on the table in this comparison. Another alternative is usually to wait for – or actively work towards – a more gradual development of the law in the desirable direction: for example, through the gradual evolution of case law or, in a federal system, through experimenting and creating support at the state levels before federal or even constitutional law is changed.

D. What are the pros and cons of each of the alternatives?

In order to make a recommendation, the various alternatives should be compared. In a *pro tanto* evaluation, most likely the pros and cons will have been restricted to a specific perspective: for instance, doctrinal or socio-legal. For fully justified recommendations, such a restriction is not possible. It should include all relevant dimensions: namely, the doctrinal and the empirical but also the normative-philosophical. Therefore, when the initial *pro tanto* evaluations are based on doctrinal research, sound recommendations require that we should also pay attention to the possible effects of a suggested legal change and to the normative-philosophical dimension. It should at least be made plausible that the change will be effective in bringing about the expected outcomes, and that the result will not be unjust. An important source for estimating possible effects could be, again, comparative legal research.⁸⁴ For example, in order to assess whether the legalisation of voluntary euthanasia is recommendable, we might want to study its effects in states

82 Smits (note 21)

83 Cf. Peter Cane, *Responsibility in Law and Morality*, 2002; Eric Tjong Tjin Tai, *Duties of Care and Ethics of Care: A Case Study in Law and Ethics*, in Van Klink & Taekema (note 3), 329–340; and Wibren van der Burg, *Law and Ethics: The Twin Disciplines*, in: Van Klink & Taekema (note 3), 175–194, respectively.

84 Smits (note 21)

that have done so, and we should also compare it with the effects of legalisation of physician-assisted suicide.

Only when all four steps are made explicit, and each step in the argument is made plausible, can the academic forum assess the argument and thus determine whether the recommendations are sound and convincing. If certain links in the chain are missing or are insufficiently supported by arguments, the forum would have to assess on the basis of insufficient information and arguments. Even if empirical research were to show that stop-and-search policies are effective in combating crime, we must be made aware of the possible sacrifices in terms of important legal values such as privacy, bodily integrity, and non-discrimination. Does it fit in a legal system with a rather restrictive approach towards granting invasive powers to the police? These aspects also have to be considered before we can make a sound recommendation to change the law.

8. Asking Too Much?

I suspect that by now many readers are keen to raise the objection that the aforementioned is asking too much. They may claim that even if it is true that we should improve the argumentative quality of publications when they contain evaluations and normative recommendations, the framework that I have suggested is unrealistic. They may argue that it requires such a broad and thorough interdisciplinary approach that providing evaluations and normative recommendations will hardly be possible,⁸⁵ and for that reason we had better stick to the familiar partial approaches that thus far have served us reasonably well.

My suggestions, if accepted, could indeed have practical consequences for legal research. The most important one is that researchers would become more methodologically reflective, and would make their methodological choices explicit and thus more open to assessment and criticism.⁸⁶ I consider that this would be an important gain. The additional implications could go two ways: legal researchers might become more cautious and they might become more ambitious.

On the one hand, if becoming more cautious, they will sometimes refrain from evaluations because they realise that the research they have done simply does not provide a sufficient basis. More frequently, they will refrain from recommendations because, as I have shown, these usually require additional arguments and a broader interdisciplinary perspective.⁸⁷ When constructing the chain of justification, they will realise that their

85 The complexity of good evaluations is a central problem in Pawson (note 19). See xvi: "Complexity sky-rockets." He concludes that it is sometimes necessary to leap. (Ibid, 30)

86 Of course, such an increased methodological reflectiveness and explicitness is not relevant only for evaluations and recommendations but may also impact legal research in general.

87 There may, of course, be other reasons for refraining from recommendations: e. g. because the recommendations would be too obvious or too controversial – or both – to different audiences. For example, in their report on the Rohingya in Myanmar, Penny Green, Thomas MacManus & Alicia de la Cour Venning, *Countdown to Annihilation: Genocide in Myanmar*, 2015, conclude that a genocidal process is underway.

research provides insufficient support for important links in the argument. For example, they have focused only on the doctrinal aspects of a statute, which is not enough to justify far-reaching recommendations of reforming it without further information about the possible effects.

Actually, it would not be a bad consequence at all if researchers were to refrain more often from making evaluations and recommendations. Doctrinal or empirical research need not always lead to evaluations, let alone to normative recommendations. The research can be of a high quality in its own right. In fact, refraining from evaluations in instances where there are insufficient grounds to justify them may improve the quality of scholarly publications. If a high-quality expository analysis is followed – and overshadowed – by superficial evaluative comments, this could have a negative impact on how readers assess the full article. In fact, I expect that it even provides some comfort to researchers in the early stages of their career when we state explicitly that evaluations and recommendations need not be an obligate part of doctoral dissertations and other publications but that, on the contrary, refraining from making them is a respectable and legitimate choice.

It may be useful to add that certain types of legal research that are not discussed in this article rarely give rise to evaluations, such as legal history or conceptual analysis. Publications in which one specific argument or counter-argument is developed may also merely allow very restricted *pro tanto* evaluations – or undermine such evaluations – but they may still be very important for the academic and public debate. This may provide a building block for a fuller evaluation. Again, this is an example of how acknowledging methodological restrictions may improve the quality of the research. Let me give an example. In the debate on opening up marriage to same sex couples, apart from appeals to religion and tradition, two arguments are frequently central to the opponents' claim. One is an empirical claim: namely, that children are harmed when they are raised by a same-sex couple; the other is a normative claim: namely, that there is an alternative, such as a registered partnership, that has precisely the same legal consequences as marriage, and therefore there is no need to change a law with regard to traditional marriage. It is, of course, completely legitimate if an author wants to debunk only one of these two arguments. She could restrict herself to an overview of empirical studies showing that the first argument is empirically false, but she could also restrict herself to arguing that the second argument should be rejected because it would leave discrimination intact at the symbolic level. Here, restriction to one discipline, either empirical or philosophical, can be preferable to trying to cover both in one article.

Researchers may also become more cautious in a different way. By reflecting on the *pro tanto* and restricted character of their evaluations, they may be able to demonstrate

As genocide is generally considered one of the worst possible crimes, this statement implies a very clear and strong evaluation. However, they provide no recommendations. I think all possible suggestions might have weakened the report. Either they would have been obvious to some and unacceptable to others (the suggestion that the Myanmar government should acknowledge that genocide is underway and should immediately stop the process) or they would have been controversial because uncertain strategic choices might be involved (the suggestion of actions by the international community).

This material is under copyright. Any use outside of the narrow boundaries of copyright law is illegal and may be prosecuted.

This applies in particular to copies, translations, microfilming as well as storage and processing in electronic systems.

© Franz Steiner Verlag, Stuttgart 2019

the limitations and strengths of their argument. If they admit openly that their evaluation is only *pro tanto*, based on a restricted doctrinal analysis or on studying the direct effects of a new statute, this makes explicit in which respects readers can rely on the outcomes. Moreover, they can indicate where fellow researchers might want to search for additional information: for example, by mentioning which relevant values and criteria have not been included in the research. They could even do so in the final section – common in other disciplines – which law publications rarely have: namely, in a section entitled ‘suggestions for further research’. In short, by clearly demarcating the methodological restrictions of their argument, researchers improve rather than weaken the standing of the publication, and they invite the forum of fellow researchers to analyse critically their chain of justification as well as to build on and to supplement it.

Justification is a matter of degree. If we try to construct a chain of justification, we may discern that some links are strong and others are decidedly weak. Acknowledging this explicitly is not a sign of weakness but a contribution to legal scholarship. Perhaps some of the chains are no more than a reference to one example, or to a few newspaper articles about the perverse effects of a new statutory rule. These sources may be the basis for weak but still initially plausible hypotheses about the effects.⁸⁸ If – and precisely because – it is indicated that these sources are relatively weak, this may also invite fellow researchers to investigate and supplement or reject the weakest links in the argument.

On the other hand, researchers may become more ambitious. The extensive framework developed in this article is gradualist. Full realisation is impossible – it is a matter of degree.⁸⁹ Justification is not a mathematical proof but a demonstration that should be as sound and convincing as possible to the academic forum. I hope that the proposed framework will challenge individual researchers as well as the larger research community – including reviewers for academic journals – in various ways to become more ambitious in achieving a higher degree of justification. For example, doctrinal research projects might be broadened by trying to find non-doctrinal evidence in the literature that may supplement the doctrinal findings.

Legal academic culture is still highly individualistic. Compared to social sciences, let alone to biomedical and natural sciences, few articles are co-authored.⁹⁰ Perhaps legal researchers should adapt to the necessity of team work; if they really want to come up with justified evaluations and recommendations, they cannot do it alone. Certainly, interdisciplinary research is not easy, but it is a road that at least a substantive group of legal researchers should take. However, it is not always necessary that we all become

88 With regard to empirical, legal, and normative claims, we should accept that we often cannot go beyond what Pawson (note 19), 81 calls “informed guesswork”. I should add that I use this phrase in a broader sense than Pawson, as he applies it only to empirical claims.

89 There is no simple standard to determine when we can be satisfied that we have produced a sound and convincing justification, also because what we should aspire to depends on the context and the forum. With regard to empirical claims, Pawson (note 19), 188 argues “we simply stop when we are satisfied that explanations are firm enough to carry the policy decision.” This basically holds not only for empirical claims but also with regard to legal, evaluative, and normative claims.

90 Epstein and King (note 6), 47–48 observe that whereas in leading political science journals almost half of the articles were co-authored, in the six US law journals they analysed only 5% were co-authored.

part of interdisciplinary teams. The argumentative framework provides a regulative ideal not so much for individual researchers but for legal research as a collective enterprise. Researchers are part of a community of scholars. They may contribute individually to a piece of the puzzle, and then others may build on that contribution. One important function of methodological sections in other disciplines is to make the research replicable, as well as show to other researchers what they do not have to do over again because it has already been solidly investigated. In legal research, replicability may not always be possible.⁹¹ Nevertheless, the readers should be able to distinguish which steps in the argument they accept because they are sound and convincing, and which steps need further backing up – or criticism.

A final remark on the claim that justification is a matter of degree. This makes it often possible to rely on shortcuts and provisionally fixed points. Let me recap three characteristics of evaluative and normative reasoning: It is addressed to a specific forum; it is plausible and defeasible; and it may rely on authoritative sources. As I argued above, quoting Rawls, arguments start from what the parties to a dispute have in common. If we can reasonably expect that the relevant forum accepts certain claims, we do not have to provide a full argument for them.⁹² The forum may accept that, in general, the most recent decisions of the Supreme Court contain the law of the United States. Therefore, if there is a recent clear judgement on an issue, for doctrinal claims we need not discuss all decisions of the Supreme Court over the last twenty years, let alone of lower courts.⁹³ A claim that the United States has accepted same-sex marriage can be supported by a simple reference to *Windsor*.⁹⁴ There are other points as well where we can rely on authority. If there is a good overview article in a prestigious peer-reviewed journal about the effects of certain types of legislation, we need not mention hundreds of other articles on the same topic. Similarly, we may refer to an authoritative exposition of legal doctrine by the leading doctrinal scholar in the field. If we appeal to democracy or the rule of law as values, we usually need not argue why they are values – although we may be required to argue why we choose a specific interpretation. Plausibility is a matter of degree, and there are no clear criteria with regard to when the degree of plausibility is high enough.⁹⁵

91 Smits (note 21), 39; Vranken (note 4), 140–141 argues that legal doctrinal research is not replicable, but should be ‘retraceable’ or ‘followable’ (in Dutch original *navolgbaar*, cf. German *nachvollziehbar*). It should be presented in such a way that the various steps in the research process can be followed critically by other researchers.

92 Perhaps we will need to do that if challenged by a critic. Legal argument is discursive, which means that we should anticipate the most important objections to our argument, and should be willing to answer any reasonable objections we did not anticipate when these are made.

93 In my view, this specific characteristic that law is a hermeneutical and argumentative discipline relying on authoritative sources, also when it includes empirical claims, is overlooked by Epstein and King (note 6). For this critique as well as other criticisms, see Goldsmith & Vermeule (note 24); for a reply not fully addressing this point, see Lee Epstein & Gary King, A Defense of Empirical Legal Scholarship: A Reply, *University of Chicago Law Review* 69 (2002), 191–209.

94 *United States v. Windsor*, 570 (2013)

95 Goldsmith & Vermeule (note 24), 164–165 argue that because of limited resources there is “a tradeoff between accuracy, on the one hand, and timeliness and relevance on the other.” In normative argument, there is a similar tradeoff between plausibility and timeliness and relevance.

Moreover, a provisional fixed point may unexpectedly become controversial: for example, widely accepted doctrinal claims may sometimes suddenly be put in doubt.

The framework developed in this paper may make it possible for researchers to present a more convincing argument to the forum, because when an argument's relative weaknesses are acknowledged explicitly, its relative strengths are highlighted. The most important result of using a stricter framework for evaluations and recommendations, therefore, should be an improvement in the quality of the research and in the quality of the evaluations and recommendations. Sketching the full framework may serve as a regulative ideal. Like all ideals it is impossible to realise; moreover, full realisation is usually undesirable. If a researcher were to attempt to meet the full framework criteria completely, she would probably never publish any evaluations. Nevertheless, it is good to keep an eye on the ideal, and to use it as a critical yardstick to determine whether we might be able to do better.

9. Conclusions and suggestions for future research

In this article, I have suggested a general argumentative framework for justifying evaluations and normative recommendations. Legal researchers should make their argument explicit and transparent, and justify the choices made for each of the links in the argument. If they are unable to do so, perhaps it is better to refrain from making evaluations and recommendations. The advice would be to do it well or not at all. Because this article is only a first step in exploring these themes, I conclude with a number of suggestions for further research. There are at least four major clusters of questions that should be addressed.

The main cluster concerns the central category of values. Are conflicting values incommensurable, or are they merely difficult to compare? Is it possible to identify and analyse the most important values? How can legal values such as doctrinal coherence and legality be operationalised, and how important are they? How can we transform efficiency and effectiveness in terms of underlying substantive values? Can we make more robust the steps from values to specific conceptions and then to specific criteria?

The second cluster concerns the differences between the context of discovery and the context of justification. The argumentative framework for justification has been presented in a linear way. However, in hermeneutic and coherentist disciplines such as doctrinal research and legal philosophy, the process of analysis is circular, going back and forth between all the relevant elements. How can we structure the process of argumentative analysis in the context of discovery? Can we develop frameworks of normative justification that will provide guidance for designing and executing legal research projects?

A third problem is that the framework is based on an unattainable ideal of full justification. To make the framework more useful in guiding actual research projects, we should try to develop legitimate partial strategies and defensible shortcuts. Some examples of partial strategies have already been mentioned, such as focusing on one specific

normative argument or making carefully restricted *pro tanto* evaluations. Another strategy regards the use of authoritative materials, from using court materials to undertaking secondary literature analysis as a supplement to doctrinal research. In order to help researchers, we should try to identify and elaborate more of these reliable shortcuts.⁹⁶

A final problem is that of incorporation. I have argued that to make justified recommendations we need to adopt the internal perspective of law. This means that the results with respect to other disciplines must be transformed so that they can fit into a legal framework with its distinct epistemic, institutional, and normative characteristics and constraints. This need for incorporation, however, raises many questions with regard to both the fact-value distinction and interdisciplinary integration.⁹⁷

We could list many more issues that need to be addressed in elaborating this argumentative framework, but the above-mentioned are the most important ones. It is my fervent hope that this article will soon be outdated because other researchers will have taken up these challenges and constructed a much richer framework.⁹⁸

WIBREN VAN DER BURG

Erasmus School of Law, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, the Netherlands; vanderburg@law.eur.nl

96 See also Péter Cserne, Facts and norms in the behavioural assumptions of law, in: Taekema, Van Klink & De Been (note 21), 105: “Here, the question is whether we know enough about a certain phenomenon to base policy on this knowledge. (...) Experience has shown various pragmatic ways of dealing with this epistemic challenge: relying on the overlapping consensus among scientists; using rules of thumb, levels of significance; or simply conducting more empirical (case, event) studies without an underlying model.” (For similar remarks see Cserne (note 14), 285.)

97 For example, see Giesen (note 14); Leeuw & Schmeets (note 14), 225–233; the special issues of the *Erasmus Law Review* 8 (2015), 2 and 3 on the incorporation problem in interdisciplinary legal research; and Taekema, Van Klink & De Been (note 21) on the relation between facts and norms in interdisciplinary legal research.

98 One further suggestion would be to analyse critically the frameworks developed in this article. This has been explorative research. I collected ideas and elements from many sources as well as from my experience in teaching research methods and in coaching Ph.D. candidates and junior researchers. I used coherentist philosophical analysis to combine these elements into coherent frameworks. For explorative research, such a heuristic and unmethodical approach is defensible, but in order to elaborate, test, and refine it, additional research might be useful. Empirical and philosophical research could provide ways to do so.