

1. Introducing empirical legal research and structure of the book*

1.1 AN APPETIZER

Characterizing empirical legal research (abbreviated as: ELR) for lawyers, legislators and regulators is not that easy. One reason is the difference between the way lawyers and ‘empiricists’ *think*.

While both modes of thinking are grounded in rigorous analysis, lawyers and empiricists often have different goals and approaches. Legal analysis places a premium on argumentation and appeals to authority, is frequently geared toward proving a particular view, is often focused on the particulars of an individual case, and is directed at reaching a definitive conclusion. In contrast, empirical analysis places a premium on observation, challenges assumptions, is oriented toward the testing of hypotheses, is usually focused on describing patterns that operate in the aggregate, and is a continuing enterprise in which new work builds on that which came before and generates even more questions for further investigation. (Lawless, Robbennolt and Ulen, 2010:10)

Although this contrast is exaggerated, there are notable differences between the two types of work.

A second reason why a characterization is difficult is that ELR makes it necessary to consider topics that students, scholars and practitioners of law are often not *acquainted* with. When you hear about empirical legal research, you probably think about data and how the data was collected. You will wonder how the (legal or societal) problem at hand was ‘translated’ into a research(able) problem. One of your colleagues may ponder the design of the study and the relationship between the ‘legal’ (normative) part and the ‘social sciences’ part. Can they be bridged? It would not be bad guesswork to think that discussions were held on what the ‘units of analysis’ are (offenders, victims, business contract partners, judges, civil servants, courts, asylum seekers, or internet service providers) and whether or not ‘we will work with a sample and use statistics’. The same goes for the researcher-jurist who is given the task to find relevant research evidence from the past. Finally, on a dog’s day afternoon, there is that young assistant visiting Wikipedia, trying to understand the role of ‘theories’ in the current project, probably wondering what is meant by the word. Purely by

accident she stumbles upon ‘legal Big Data’ and ‘disruptive legal technologies’ (and sends a message to her colleagues that she’s a bit lost).

These topics are all relevant and ‘in-need-of-being-known’ for anyone planning to *do* ELR, to *commission* ELR and/or to *use* (and *evaluate*) it. However, they are also fragmented and miss an overarching characterization of what *empirical legal research* basically is.

This guidance book is dedicated to prepare and present such an overarching perspective. Its goal is to give guidance to students, scholars and practitioners, be they lawyers, legislators, regulators, policy makers, commissioners of research, prosecutors or judges. A complete ‘menu’ of what empirical legal research entails will therefore be offered.

Good dinners start, according to the norms of *haute cuisine*, with an *amuse bouche*. Several of these are now served as well as the concept of the ‘empirical cycle’.

1.1.1 ‘Law in the Books versus Law in Action’

Empirical legal research (ELR) covers all major fields of law, as the *Oxford Handbook of Empirical Legal Research* (Cane and Kritzer, 2010) shows. The chapters include family law and penal law, access to (civil) justice, evidence law, contract law and international law, but also police activities, naming and shaming, regulatory impact assessments, the role of law and regulation with regard to topics such as bankruptcy and financial markets and consumer protection.

Empirical legal research, however, is *not* primarily or only interested in *laws* (or any other legal arrangement) *in the books*, but in *law(s) in action*, to quote Pound:

If we look closely, distinctions between law in the books and law in action, between the rules that purport to govern the relations of man and man and those that in fact govern them, will appear and it will be found that today also the distinction between legal theory and judicial administration is often a very real and a very deep one. (Pound, 1910: 15)

Llewellyn (1930), a few decades later, distinguished between ‘paper rules’ and ‘real rules’ or ‘working rules’. Laws and rules can exist everywhere, although they are often ‘seen as a body of rules that . . . are determined and enforced by the state’ (Shavell, 2002: 227).

In this book we broaden this concept in two ways. First, because there are more organizations than states producing rules, regulations and related interventions like ‘private regulatory regimes’ (Aviram, 2003), ‘legislative Marktinterventionen’ (Hosemann, 2014: 45), publicly declared (corporate) codes (as binding promises) (Beckers, 2015) and international treaties and

protocols. Second, because as Howarth (2013) highlights, the legal world designs ‘social structures and devices [like] contracts, companies, trusts, constitutions, and statutes . . . A successful contract is one that gets the job done of facilitating a deal between the parties. A contract that results in litigation is unsuccessful – it has failed to do its job’.¹

In this book we use different terms to describe the object of empirical legal research: (*legal*) *arrangements, devices, interventions*, but we will also refer to *laws, legislation* and *regulation*. All these ‘legal arrangements’ aim at influencing individuals and organizations, codify and/or regulate (to some extent) their functioning, create restrictions and opportunities for them and can be enforced with various degrees of involvement of the state.² Westerman (2011: 106–107) adds that ‘many rules nowadays prescribe in a fairly direct manner the goals that should be attained’ and calls this the ‘ongoing instrumentalisation [that] has led to different demands on both the legislator and judge’. Verbruggen (2014: 79) sees ‘commercial contracts [becoming] increasingly important vehicles for the implementation and enforcement of safety, social and sustainability standards in transnational supply chains’.

ELR thus addresses developments and actions in the ‘real (social) world’³ as relating to legal arrangements, either to influence this world, to facilitate it, or to legalize what has been the ‘usual’ practice.

One of the backgrounds of ELR has been the critique, articulated decades ago, on legal formalism; some even referred to the construction of a ‘heaven of legal concepts’ far removed from social reality, and disregarded the ways in which law is produced by and operates within society (Cohen, 1935). Although this statement is an exaggeration and causes unnecessary skepticism about *any* role formal laws and rules play in reality, differences between the ‘legal reality’ and the ‘social (empirical) reality’ cannot be denied.⁴ The distinction between doctrinal (legal) research and non-doctrinal studies is related to this point.

Doctrinal research asks what the law is on a particular issue. It is concerned with an analysis of the legal doctrine and how it has been developed and applied (McConville and Chui, 2007). This approach, according to Hutchinson (2013: 9) ‘lies at the heart of any lawyer’s task because it is the research process used to identify, analyze and synthesize the content of the law . . . Doctrine has been defined as a synthesis of rules, principles, norms, interpretative guidelines and values, which explains, makes coherent or justifies a segment of the law as part of a larger system of law’. McCrudden (2006: 634) adds that ‘doctrinal analysis takes [the] form, in which the writer attempts to argue that this or that is the “best” solution to a particular problem, “best” meaning having the best fit with which already exists’. Non-doctrinal research is legal research that employs methods taken from

other disciplines to generate empirical data to answer other research questions. It can focus on a problem, its causes and consequences, a policy, a law or any other legal arrangement while institutions and organizations can also be studied.

ELR focuses on ‘empirical’. In line with the editors of *The Oxford Handbook on Empirical Legal Research*, this ‘involves the systematic collection of information (“data”) and its analysis according to some generally accepted method. Of central importance is the systematic nature of the process, both of collecting and analysing the information’ (Cane and Kritzer, 2010: 4–5). Epstein and Martin (2014: 3) use a more informal terminology: data is ‘just a term for facts about the world’, sometimes numerical (or quantitative), sometimes non-numerical (or qualitative).

How does empirical legal research work look in practice? A few examples: Smits (2011) studied mandatory rights of withdrawal in *consumer contract law* and what consumers think about these rights. Withdrawal rights allow the consumer to terminate the contract within a set ‘cooling-off period’. Smits made a comparison between statutory withdrawal rights in Europe and in the United States. He presented results of a ‘modest survey’ of the voluntary use of withdrawal rights in general conditions of retailers. He also studied the usefulness of mandatory withdrawal rights.

Ruiter et al (2011: 135) investigated the (*penal*) *sentencing preferences* of the general public versus those of judges in the Netherlands. ‘Given the opportunity, how would the Dutch public sentence perpetrators of different types of crime? To what extent are these verdicts related to characteristics of the criminal act (e.g. offender characteristics, type of crime, victim characteristics)? Does the verdict depend on characteristics of the citizen who issues a sentence (e.g. young/old, male/female)? And does information about sentencing options (i.e. the costs involved and rates of recidivism) affect verdicts decided by members of the general population?’

Increasingly the impact of regulation and other (legal) arrangements is studied from an experimental or quasi-experimental perspective: individuals, groups or organizations are compared, over a certain period, while some get a ‘treatment’ (the (legal) intervention, program, sanction, piece of legislation, contract) and others do not. Interventions like Hot Spots Policing (Braga et al, 2012) are studied along these lines, but there are many other examples in the field of private and administrative law.

ELR also covers studies of a different nature: studies describing, analysing and comparing ways in which – as an example – nation states safeguard the *rule of law* or *freedom rights*. The ‘World of Justice’ project (Agrast et al, 2014: 2) presents a set of empirical indicators on the rule of law from

the perspective of the 'ordinary person'. It examines practical situations in which a rule of law deficit may affect the daily lives of those interviewed. The project provides data on 10 dimensions, such as order and security, absence of corruption, regulatory enforcement and access to justice. McMahon (2012) is one of the authors behind the *index of human freedom* which combines country data on the freedom of movement, freedom of expression, and freedom of relationships. It also covers data on government's threat to freedom (the occurrence of political imprisonment) and society's threat to freedom (like human trafficking, homicide and female genital mutilation).

1.1.2 The Functioning of Organizations and Institutions in the Legal Field

ELR is not only focused on *laws and regulations (in action)*, but also on the *functioning of organizations and institutions in the legal field*. What they do, how they do it, what the consequences are of their work and what they cost are topics of interest. The legal field is characterized by large numbers of organizations and institutions. National and international courts, prosecutors, prisons and probation services, but also ombudsman institutions, human rights organizations, bailiffs and notaries, and their professional organizations, including what Jansen and Michaels (2007) call private law beyond the state (multinational companies creating systems of rules and imposing them on their suppliers and customers; examples are standardization and accreditation organizations and transnational governance networks).

Van Dijk et al (2009) studied one such institution, the *Netherlands Ombudsman*. They used the *Ombudsman's complaints database*, which contains information from over 140,000 dossiers covering 25 years. They applied a mix of conventional database techniques and a data mining algorithm to scan and map this database and were able to detect paths in the handling of complaints that were not known before.

Posner and de Figueiredo (2004) present a second, rather different example. They studied the International Court of Justice (ICJ) which has jurisdiction over disputes between nations. Its defenders argue that the ICJ decides cases impartially and confers legitimacy on the international legal system. Its critics argue that the members of the ICJ vote in the interests of the states that appoint them. Prior empirical scholarship is 'ambiguous', as Posner and de Figueiredo (2004) showed, which urged them to undertake a new empirical investigation. They studied theories on judging behavior and tested the charge of bias using statistical methods on voting behavior of Court members.

1.1.3 Evaluating Laws, Regulations and Other Legal Arrangements like Interventions and Sanctions

Doing ELR often implies asking and answering *evaluative questions*. Cummings (2013: 186) refers to the ‘challenges of evaluation’ that empirical legal researchers face. ‘How to study the impact of law is a question that has confronted scholars since the 1960s. There are questions about which types of impacts to study and the proper methodology for doing so’. Hage (2010: 6) refers to ‘evaluative legal science which can take the shape of passing value judgments on actual or hypothetical (proposed) regulations or of the selection of the “best” regulation from a set of alternatives’.

Evaluation is the more so important because, as Howarth (2013: 15; 17; 67ff) argued, lawyers in fact are (transaction-costs-) engineers: they make devices for others that try to reduce or prevent transaction costs, conflicts and other problems (Gilson, 1984).⁵ Devices are ‘contracts, conveyances, wills, trusts, regulations, statutes and constitutions, and companies’ (i.e. their legal structure).⁶ However, between ambition and reality can be discrepancies, and evaluators are ‘designed’ to monitor and explain them.

Evaluators also study the predictions (of public and semi-public organizations) about the consequences of policies, rules, legislation and other ‘tools of governments’ (like information campaigns and incentives), sometimes before they are even implemented. This is known as *ex ante evaluation* or ‘prospective evaluation’. The European Union refers to these activities as *regulatory impact assessments*.⁷ A US GAO (1996) study trying to find out what the impact on teenage pregnancy would be if law A or law B would be implemented, is a (classic) example.

Evaluations also address the way in which legal arrangements are implemented and what difficulties are experienced during this process (*process evaluations* or *implementation evaluations*). A related approach is known as *compliance (or regularity) auditing*. Here the empirical investigation describes the extent to which natural and corporate actors behave in compliance with protocols, rules and regulations.

And there are studies (often *ex post*) on the *effects, consequences or impacts* of legal arrangements and interventions on the behavior of persons and organizations (*impact or effectiveness evaluations*). A rather old example is Aubert’s study done in the early 1950s. He wanted to find out the ‘extent to which behavior (of housemaids) conformed to the rules laid down in the (Norwegian) Law on Housemaids of 1948’. The purpose of this law was to protect the interests of domestic help. Aubert used a sample of some 200 housewives and 200 housemaids in Oslo and interviewed them about their conduct, level of knowledge, and attitudes and motives, in so far as these related to the law. ‘It has to be concluded that the law was, at least for

some years, ineffective in the sense that actual conditions of work remained at variance with the norms laid down' (Aubert, 1969:121).⁸

Interventions in the world of penal law not only refer to regulations and sanctions (for example with regard to the role they play in reducing recidivism), but also include (cognitive-behavioral) programs used by prisons and probation organizations to alter attitudes and behavior of (re)offenders. Examples are ETS ('Enhanced Thinking Skills') and ART ('Aggression Regulation/Replacement Training'), which are regularly evaluated (Cornet et al, 2014).

Since the introduction of the internet, laws, regulations and (other) policies are confronted with and directed at the digital world. Evaluating the impact of (digital) sanctions to help prevent or reduce digital piracy (illegal copying of music, movies, 3-D printing designs and e-books) is only one example; others deal with evaluating the governance (structure) of the internet, e-law enforcement, circumvention policies and digital surveillance activities (Leeuw and Leeuw, 2012).

Evidence-based regulation, wherein results from social science research including behavioral economics are used to help law makers draft regulations that have a (fair) chance of realizing their set goals, is a blossoming field. Sunstein (2013: 1) states:

In recent years, social scientists have been incorporating empirical findings about human behavior into economic models. These findings offer important insights for thinking about regulation and its likely consequences. They also offer some suggestions about the appropriate design of effective, low-cost, choice-preserving approaches to regulatory problems, including disclosure requirements, default rules, and simplification.

Parts of this approach are the so-called design studies and pilots. Alternative dispute resolution activities, for example, are developed and tested before being implemented at a broader scale. The Hague Institute for the Internationalization of Law⁹ shares examples of how research is related to modernizing justice.

Throughout this book, examples of legal evaluations are on the menu, including the designs that are used and why, what the role of 'intervention theories' are, how data is collected and analyzed and how evidence from legal evaluations can be visualized in a modern way.

1.1.4 Big Data, Technology and the Law: Legal Predictions, Machine Learning and Computational Legal Studies¹⁰

Big Data and *technology* (artificial intelligence, machine learning and the internet) are strongly related. Big Data consist of (call) logs, mobile-banking

transactions, online user-generated content such as blog posts, social media, sensor apps, online (Google) searches, satellite images and the data one holds in emails, blogs, intranets, extranets etc.¹¹ The *Internet of Things* refers to the ability of everyday objects to connect to the internet, allowing these devices to gather, send and receive data. Examples include wearable technology, such as watches and fitness bands or smartphones but also thermostats that know when we are home and cars that monitor our driving habits (O'Neill, 2015). The more (historical) paper documents like wills, contracts, treaties and other legal case materials, as well as books and newspapers, become digitized, the larger 'legal Big Data' becomes.

Susskind (2008; 2013) refers to these and other technological developments as *disruptive* (for the legal world). Examples are *automated document assembly* (disruptive, because widespread use can greatly reduce the time that lawyers expend on document drafting and production); *relentless connectivity* (he uses this term to describe how the use of hand-held devices with wireless broadband access, powerful video, high processing speed and nearly endless storage capacity will create expectations among clients for 24/7 lawyer availability) and *online legal guidance* (systems that provide legal diagnoses, generate legal documents, assist in legal audits and provide legal updates).¹² With regard to Big Data, Süsskind (2013: 48–9) suggests that 'in due course they will be of profound significance for legal practice and scholarship'. A question like 'what legal issues are troubling communities can very easily be answered, while by collecting and analyzing huge bodies of commercial contracts and exchanges of emails, we might gain insight into the greatest legal risks that specific sectors face'. Online dispute resolution (ODR) is materializing¹³ while social media can help in finding out which public perceptions exist about magistrates.

Also for ELR Big Data are important. One reason is that (a new form of) *legal analytics*, including *legal prediction* is possible. McGinnis and Pearce (2014: 3052) put it as follows:

Law, with its massive amounts of data from case law, briefs, and other documents, is conducive to machine data mining (machine learning)¹⁴ that is the foundation of legal analytics. Legal data include fact patterns, precedents, and case outcomes. For instance, one form of legal analytics would use fact patterns and precedents to predict a case's outcome, thereby better equipping lawyers to assess the likely result of litigation.

Nelson and Simek (2013:1) refer to 'using Big Data to *evaluate* (law) firms'. Another application is that Big Data will stimulate *transparency* of the legal world. Katz and Bommarito (2013: 3–4) studied how complex legal rules are (in the USA). Big Data may also facilitate *legal practice/legal aid*. Empirical researchers doing pattern recognition in legal advice, complaints

and electronic transcripts of court procedures and know-how can play an important role in aid processes. In criminal investigations, tax fraud detection and detecting money laundering already play such a role. Finally, Big Data can deliver *new evidence* in court. Several American examples indicate that big data collected and analyzed from public data sets can be admitted as evidence. Although the legal sector has always been a 'data-driven industry', until recently all that data stayed offline (i.e. on paper).

Some of these examples come from the USA, and indeed, the Big Data movement in law seems to be more developed than in Europe. However, it can be expected that, as has happened in other fields (marketing and business analytics), the European legal world will soon catch up. One reason is that digitization does not stop at the frontier. Another reason is that while in the justice domain many organizations, such as the police, public prosecution, courts and prisons have their *own* data and information systems, society expects empirical insights *throughout this domain* on *overarching* items like judiciary efficiency, case load, elapsed time and privacy risks.¹⁵ These insights can only be produced when judicial data space systems are available, covering the diverse organizations' own data and making domain-wide analysis of the 'enterprise of law', including legal logistics possible.

1.1.5 Explaining what is Happening: the Role of Theories

Observing that legal arrangements make a difference (or do not) is something; to explain what has happened is something else. *Explanatory questions* have to be asked, like what are the underlying causes that lead some legal arrangements to trigger behavior(al changes), while others are falling dead on the ground? An example is MaCaulay's (1963) study on (non-) contractual relations in businesses. He found that in the USA, contract law is often ignored in transactions, which he tried to explain by using insights from sociological and behavioral theory. Since this path-breaking study, a research program has developed, investigating under which conditions 'contractual behavior' occurs (see Chapter 10).

Explaining why wrongful convictions occur is a second example. Technical mistakes during police investigations and incorrect eyewitnesses are one aspect of such an explanation, but cognitive biases of officials like tunnel vision, yeah-yeah-saying, hindsight bias, and the fundamental attribution error¹⁶ also help explain the occurrence of erroneous convictions.¹⁷

A third example has to do with explaining a (lack of) compliance with rules and regulations. It is known that some people and organizations are more focused on compliance than others. A Dutch study showed that with regard to the introduction of cognitive-behavioral interventions like

Enhanced Thinking Skills (ETS) into the Dutch penitentiary system, around half of the interventions were not implemented in accordance with procedures and protocols agreed upon (Nas et al, 2011). Insights from *public choice theory* help to explain this. One such insight is that there is competition inside bureaucracies about (the number of) clients to be ‘treated’ (‘by me’). In particular, when numbers of clients are small (or decreasing), bureau-politics and bureaucratic firefighting pop up as factors causing implementation problems¹⁸ (while the bill is paid by taxpayers).

To explain why individuals differ in levels of compliance with the law, *rational choice theory*, sometimes in connection with *biosocial theories*, is used.

Bio-social theories – in this example – point to differences in cognitive and emotional capacity (‘agency’) of persons to perceive and experience threats of being caught. Fearlessness theory, low arousal theory and other theories from endocrinology, neurosciences and genetics are important points of reference (Raine, 2013). These insights are also used to explain what the impact of behavior modification programs is on (re)offenders (Cornet et al, 2014).

We have outlined five important characteristics of ELR, and used the word ‘empirical’ rather often. What is it and what is the empirical cycle?

1.2 THE EMPIRICAL CYCLE

What is ‘empirical’? The editors of the *Oxford Handbook on Empirical Legal Research* describe ‘empirical’ as ‘the systematic collection of information (“data”) and its analysis according to some generally accepted method’ (Cane and Kritzer, 2010: 4–5; Lawless, Robbennolt and Ulen, 2010: 7). ‘Systematic data-collection and analysis’ and ‘accepted methods’ are important elements. What this entails will be discussed in the following chapters. However, we emphasize that ELR is much more than data-collection and analysis. It includes research problems, theories, systematic literature reviews, research designs and knowledge transfer.

A *second characteristic* is that the empirical cycle *includes both quantitative research and qualitative research*. This may seem an entirely superfluous remark, but it is not. Some of you remember or have heard about ‘paradigm wars’¹⁹ between approaches (in the latter part of the twentieth century). Although these times are behind us, there remain epistemological and methodological differences. While quantitative research often employs experimental designs and quantitative measures to test hypotheses and emphasizes the measurement and analysis of causal relationships between variables, qualitative research uses a more naturalistic approach that seeks

to understand phenomena in ‘real world settings’. Unlike quantitative researchers who seek causal determination, prediction and generalization of findings, qualitative researchers seek illumination, understanding and extrapolation to similar situations (Golafshani, 2003). Computer-assisted qualitative research methods have reduced the differences between ‘quali’ and ‘quanti’, while the introduction of Big Data, including sensor data, will further reduce ‘paradigmatic firefighting’.²⁰

A third characteristic is the *cyclical nature* of empirical research. Empirical research starts with a question (‘research problem’), passes through several stages of activities like data collection and analysis, formulating an answer to the problem and often starts again in another cycle addressing a related research question.

The cycle starts with → (research) problems: In every discipline and specialty, research starts with (*research*) *problems* (Popper, 1972; Ultee, 1976; 2001; Bunge, 1997). The same is true for ELR. These problems find their origin in different things:

- in prior research (‘unsolved scientific problems’);
- in scientific debates (contradictions between findings and theories, mixed results, different interpretations, known unknowns and unknown unknowns);
- in trends, developments and problems in the physical, social, legal and virtual world.

Examples of research problems addressed in the *Oxford Handbook on Empirical Legal Research* and in recent empirical legal research are the following:²¹

- What are the possible reasons for not holding executives and oversight officials internally liable for the damage caused by the improper performance of duties by agencies in cases where there is sufficient ground to do so?²²
- How can claiming behavior be explained and what are findings from empirical research on naming, blaming and claiming?
- To what extent is Automatic Number Plate Recognition (ANPR) technology an efficient instrument to improve investigations by police and prosecutors? What are (unintended) side effects of implementing this technology in the Netherlands?
- What are assumptions about human behavior underlying the decisions made in refugee status determinations by judges in the UK?
- Has federal antidiscrimination law in the USA been effective in moving women and minorities into management?

After the research problem(s) have been articulated → ‘theory’ marches in (most of the time): A theory ‘is simply a reasonable and precise answer to

the research question' (Epstein and Martin, 2010: 906). Although there are several types of theories, in this book we discuss two of them: *formal, aka explanatory theories* and *intervention, aka program theories*. Formal (explanatory) theories often have been developed by the academic community and have a label or a name. Examples are public choice theory (on decision-making mechanisms in publicly financed organizations), cognitive dissonance theory (on balancing cognitions, attitudes and behavior) and incentive theory (on different types of incentives and prompts). Intervention or program theories are sets of assumptions underlying (public/private) interventions like laws and other regulations, contracts, covenants, verdicts, levies, subsidies, awareness raising campaigns and naming and shaming strategies. Often these assumptions are not (fully) articulated by practitioners, policy makers and other users and producers. The assumptions imply the answer to the question why they believe that their interventions, operations or verdicts are capable of reaching certain goals or at least contributing to them.

Related to the theoretical work, researchers want to find, systematize and synthesize results from earlier research → research reviews and synthesis marches in: Researchers want to know what has already been published on 'their' problem. In earlier times, surveys of the literature were based on suggestions from friends, colleagues, librarians and journals ('snowball sampling'). Over the last two to three decades a more rigorous type of literature study has been developed, known as the *systematic research review* (and *synthesis*).²³ Based on a broad (on-line) coverage of literature through search engines like Web of Science, other (meta-)search engines and websites from publishers, the field is scanned and mapped, while inclusion and exclusion criteria (which studies will be used and which not) are applied (based on protocols). One of the goals is to distinguish between empirical and non-empirical studies, and to take into account the robustness of the research designs and (data collection and analysis) methods (including statistics). These '[systematic review] techniques have been developed to weigh studies against each other in terms of the confidence with which their findings can be accepted and to combine data from different studies to reach overall judgments' (Gomm, 2008: 345; see also Gough et al, 2012).

Often this review work addresses the research problem directly; if the problem concerns the impacts of a restorative justice program in Canada, then a systematic review of empirical studies of (similar and other) restorative justice programs is very probably available. Another approach is to carry out a systematic search and review of earlier work focusing on the *mechanisms* underlying legal arrangements, interventions and policies (Farnsworth, 2007). If the research problem is to find out what the impact is of legal interventions to curb digital piracy, and there is a theory

specifying which behavioral mechanisms may explain (and change) this behavior, information from empirical studies about *these mechanisms* can be retrieved (Leeuw and Leeuw, 2012). However, research reviews may not be available, or the results may be contradictory or apply to fundamentally different contexts.

Then: the decision is to continue the project by doing new (primary) research which necessitates that the research design appears on the agenda: To collect *empirical information* in a methodologically robust way, decisions have to be made on the *research design*. The function of such a design is to ensure that the evidence obtained enables us to answer the initial question as unambiguously as possible. Research designs structure the way in which a study is constructed. Examples of designs are the experiment, the quasi-experiment, the case study and the longitudinal observational study. Designs are *not* the same as methods to collect data like interviews, focus groups or questionnaires.

Next → the choice of data collection methods marches in: There is wide variety of data collection methods available: questionnaires, item lists (for interviews), observation schemes and ‘reading lenses’ to categorize and summarize documents (like reports, contracts, proceedings of meetings, e-mails, blogs, websites), but also videos (YouTube) and pictures, psychological, physiological and neuro-scientific tests, sensors and digital ‘apps’. Both quantitative and qualitative methods are part of this toolkit as well as multi-method approaches (Nielsen, 2010). Deciding about who the units of analysis are is important too.

Sometimes data are already structured and stored and (almost) ‘ready for use’: databases and data warehouses are the places to be. Stanford University’s Faculty of Law has the Intellectual Property Litigation Clearinghouse, which provides real-time data on intellectual property litigation. The Research Institute of the Dutch Ministry of Security and Justice (<http://www.wodc.nl>) has databases covering penal law, administrative law and civil law.

Collecting data is one thing, analyzing them is another thing → now data analysis is on the menu:

How do empirical legal scholars perform this task? The answer depends in no small part on their goals. If the goal is to summarize the data they have collected, then some simple statistical measures might suffice. For the vast majority of ELR projects, however, making inferences (using facts we know to learn about facts we do not know) is the goal. (Epstein and Martin, 2010: 912–13)

Then more complex statistical (descriptive and causal) activities, with the help of software, are carried out. However, not every ELR study is quantitative in nature. More qualitative data can be collected through interviews,

focus groups or participatory observation. Also, documents like wills, contracts, verdicts, jurisprudence and policy papers can be analyzed. Special methods including software to analyze this kind of data are available.

The penultimate step → presenting the results of ELR and answering the research question(s): This was usually done through a book or report, a paper or a fact sheet. Nowadays, in order to transfer findings to relevant audiences in a more modern way, use is made of seminars and webinars, blogs, apps and round table meetings with stakeholders, colleagues and commissioners. Data-visualization has become important too. What is also at stake is ‘the problem of adding normative value to facts: no legal obligation follows from empirical facts’ (Lepsius, 2005). This topic is known as the fact-value-dichotomy or the gap problem (between empirics and normativity). ELR folks should not only be capable of understanding this problem but also detecting and handling it. The same is true for normative-doctrinal researchers (Leeuw, 2015).

*Then the process may start again → a new (related) research problem is formulated and a new research cycle starts . . .*²⁴: It may start again because the investigators have not been able to answer the research problem (in full) due to (methodological) difficulties, data restrictions or other causes. The research cycle can also start again (on a ‘higher’ level) when the research problem is answered, but leads to *new* questions.

Graphically speaking this is the *empirical research cycle* as presented in Figure 1.1.²⁵

As the examples presented show, ELR has breadth, depth and width. *Breadth*, because it covers all legal fields and specialties. *Depth*, because it studies these fields by combining what is happening with laws and regulations ‘in the books’, and ‘in action’, by applying different approaches, designs and methods of empirical research. And *width*, because it does not look into research topics from one perspective (like compliance audits do) but addresses them *side to side*, i.e. at its full extent.

Before we move on, a few words of caution. Is this circle, with its six steps, a ‘mechanical process from which the analyst can never deviate’ (Epstein and Martin 2014: 19)? They compare the analyst/researcher who strictly follows such an approach with a ‘baker who must follow every step in the recipe or else the cake will fall’.²⁶ For them, researchers ‘should have the flexibility of mind to deviate, to overturn old ways of looking at the world, revise their blueprints as necessary and collect more (or different) data than they might have intended’. We agree with the idea that researchers have the intellectual and professional freedom to deviate, *but at the same time research is not an ‘anything goes’ process, a free floating ‘pick and choose-thing’*. Formulating low-quality research problems, leaving theory outside the investigation, not knowing right from wrong when discussing

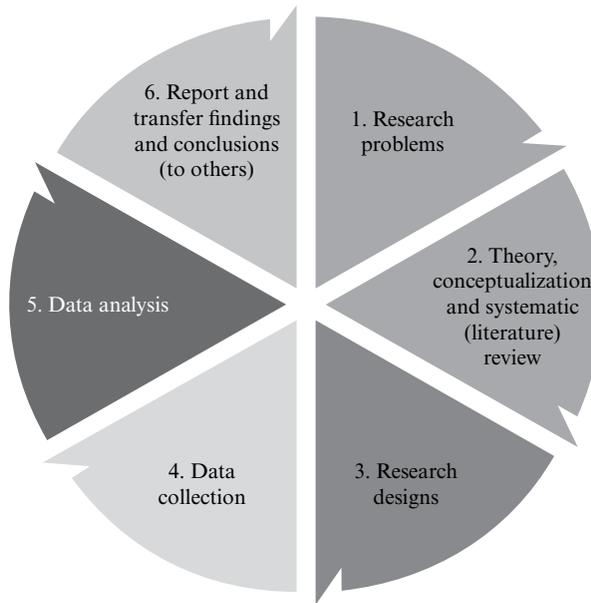


Figure 1.1 *The empirical cycle*

and implementing research designs, forgetting about validity when collecting and working with data and neglecting the fact-value dichotomy, is *not* the way to go.²⁷

Although Ramsey and Keller's (2006) list of 11 tips is designed for judges and lawyers when *using* empirical research, insights as to what is commendable when *doing* empirical work in a legal environment can easily be deduced from it. The authors stress the relevance of an adequate research problem and hypotheses, the importance of how concepts are operationalized and measured, the causality problem (and which research designs are capable of addressing it) and several others.

My suggestion to the reader is: use and follow the different steps, the related methodological 'rules of thumbs' and criteria, but do so in a *smart way, preventing it from becoming a tick & flick-box or iron cage.*

1.3 STRUCTURE OF THE BOOK

Chapter 2 starts with looking back. It gives a *concise contemporary history* of ELR and discusses its 'intellectual roots'.

BOX 1.1 ON THE QUALITY OF EMPIRICAL LEGAL RESEARCH

For the USA, Epstein and King (2002) studied the problem of (methodological) quality of (US) empirical legal research. Initially they read and analyzed all 231 articles published in American law reviews between 1990 and 2000 that had the word 'empirical' in their title. Their purpose was to establish what current practices in the US law reviews were and what were the (methodological) problems. Later they added all empirical papers from six (top) US law reviews (Chicago, Columbia, Harvard, NYU, Stanford and Yale) published between 1995 and 2000, the 50 most-cited articles (according to the Legal Scholarship Network) that were written by legal academics and appeared in the law reviews and studies in four peer-reviewed law journals (like the *Journal of Law & Economics* and the *Journal of Legal Studies*). Finally, they used 'a much more informal approach; namely, by reading widely through law reviews, following citations, and reading further'. The results were 'discouraging', which led them to formulate guidelines on how to do empirical legal research. Most of them directly boil down to the dimensions of our empirical cycle outlined above, while a few are broader in nature.

In the Netherlands, Smits (2009) has criticized the state of the art of (empirical) legal research, as did the Report of the Evaluatiecommissie Rechtswetenschappelijk Onderzoek. Tijssen (2009: 197, 206) analyzed some 100 PhDs in legal research: 'The (Phd) research offered an evaluative test of the methodological justification of the legal research presented against general methodological norms'. His conclusion was that 'in terms of verifiability, legal doctoral theses do little in the way of explaining and justifying the sources selected and the methods used, yet an increasing number of doctoral theses include an explanation and justification of the research problem'. For the field of human rights research, Coomans, Grunfeld and Kamminga (2008) are of the opinion that 'in academic human rights research, especially legal human rights research, little attention tends to be devoted to questions of methodology'.

Next, the focus shifts to describing the *different aspects (steps or dimensions)* of the empirical cycle.

Chapter 3 focuses on (typologies of) *research problems*, including *suggestions on how to develop adequate formulations of problems*.

In Chapter 4 the role of *theories* is highlighted. *Two types* of theories are discussed (explanatory and intervention theories), what their *ingredients* are (mechanisms), how theories can be found and/or articulated, including examples.

Systematic research reviews and synthesis studies as ways to capture *earlier research* are discussed in Chapter 5. Types of systematic reviews are presented as well as information how to find them and how to do them.

Chapter 6 brings types of *research designs* to the menu. Criteria are presented for selecting and assessing design(s) that fit the type of research

problems, including methodological characteristics of designs as well as examples.

Chapter 7 summarizes *data collection methods and instruments*, both *quantitative and qualitative*, including *ethical considerations*. Three types of data are discussed: *register-based (and stored) data*, *data collected through primary research and Big Data*. Reliability and validity of data (collection) are discussed as well as other criteria, including ethical.

Chapter 8 gives information, suggestions and examples on *quantitative and qualitative data-analysis*, including *statistics, Big Data analysis and mixed methods*. A large part of this chapter is structured around several empirical investigations dealing with elections in OSCE regions. The chapter also describes data-visualization, including some of its tools.

Chapter 9 discusses two issues on how to handle the transfer and dissemination of findings from empirical research. The *first* is how to help the *diffusion/transfer of the evidence to diverse audiences*, in such a way that the results are known. The *second* issue discusses the (complex) relationship between empirical legal research and normative aspects, including legal practice and scholarship.

Chapter 10 asks two final questions. *Is an empirical revolution in law taking place? And: is there growth of knowledge in the ELR field?*

ENDNOTES

- * Thanks to Jaap Hage, Miet Vanderhallen, Willemien de Jongste, Frank Willemsen, Cor Hermans, Jack Tomlin, Katja Zimmerman, Honours Track master students from Maastricht University and members of the Maastricht Law and Philosophy Platform for their comments and suggestions.
- 1. Howarth compares the work of lawyers with engineering. He was not the first, as Pound (1942: 64) saw law as 'social engineering designed to eliminate friction and waste in the satisfaction of unlimited human interests and demands out of a limited store of goods in existence'. See also McManaman (1958: 16ff).
- 2. Hadfield (2001), Schwarcz (2002) and Ellickson (1991) point to the lengthy historical precedence of this type of legal arrangements. Hadfield (2001: 40, 41) shows that 'from the middle ages to infant digital age, there are examples of law developed and administered by private entities with various degrees of state involvement'.
- 3. This includes virtual reality, augmented reality and, in the not too distant future, inter-reality (where virtual and off-line reality merge).
- 4. Schauer (2013) presented a balanced view on differences between 'law in books/ in action', including types of rules. Social and behavioral researchers are well-versed in understanding the role formal rules/legislation play, as they influence incentives and other mechanisms that underlie behavioral choices and outcomes.
- 5. Susskind (2008) sees lawyers as 'knowledge engineers', who work as knowledge brokers together with expert systems. See also Howarth (2013: 20–21).
- 6. In the 1950s (in the USA) a new discipline: 'economics was suggested, the study of good order and workable arrangements' (Fuller, 1953: 477). See also Polish sociologists Podgorecki and Petrazycki, who worked in line with social engineering.

7. In the evaluation literature 'impact assessment' is usually done *ex post*, to find out what the effects ('impact') of policies and programs have been.
8. Seen from a contemporary perspective, the design that was used can be criticized. See Chapter 6 for more on research designs.
9. <http://www.hiil.org/search?form=search&q=trends>, accessed 4 July 2015, describes the many 'modernizing' activities of this institute, including trend and future studies.
10. Sometimes the concept of money-ball legislation/regulation is mentioned. It goes back to Michael Lewis' 2003 book *Moneyball*. 'Lewis tells the story of Billy Beane, the general manager of the Oakland Athletics baseball team. With the help of his statistics-obsessed assistant, Paul DePodesta, Beane brought the once lowly Athletics into the top tier of baseball teams, and wound up transforming professional baseball, by substituting empirical data for long-standing dogmas, intuition, and anecdote-driven judgments' (Glaeser and Sunstein, 2014: 1). As Moneyball legislation largely uses statistical data and findings from cost-benefit studies and policy analysis, and – as far as we know – has not yet worked with 'Big Data', we do not discuss it here (although the underlying idea, that data are important when designing and implementing regulation, is the same).
11. Nelson and Simek (2013: 2) give examples ranging from search and rescue persons and law enforcement agencies working with drones to 'the Occupy Wall Street protesters (who also) had their own drone' (The Occucopter).
12. See a selection of the results from the 2010 American Bar Association (ABA) Legal Technology Survey Report, http://www.americanbar.org/groups/departments_offices/legal_technology_resources/resources/articles/youraba0610.html, accessed 17 July 2015. During European Law and Technology conferences disruptive technologies are also discussed (<http://www.prweb.com/releases/2012/8/prweb9779900.htm>, accessed 16 July 2015).
13. To give an example: LetsGetDivorced.com is an online self-directed legal document assembly service for people looking to file a simple uncontested divorce, where both spouses agree on all issues of their divorce, such as equitable distribution of their marital property and marital debts, spousal support, child custody, visitation and child support.
14. Machine learning is a type of artificial intelligence that provides computers with the ability to learn without being explicitly programmed. Machine learning programs detect patterns in data and adjust program actions accordingly. Examples range from speech recognition, to web searches, the understanding of the human genome and robot control.
15. The elapsed time of a case is the period that is required to handle a case that pertains to a suspect or convict. Long elapsed times may be interpreted as delays in a criminal justice system, which in turn may lead to 'justice delayed, justice denied'.
16. Other examples are confirmation bias (searching for confirmation and verification), the Texas sharpshooter fallacy (perceiving patterns in random events that do not exist) and the Lake Wobegon mechanism (observed among drivers, CEO's, coaches, radio show hosts and others). It was dubbed the (fictional) 'Lake Wobegon Effect' in tribute to the fictional radio comedy community of Lake Wobegon, where all the children are above average.
17. Lawless, Robbennolt and Ulen (2010: 14) use this phenomenon to stress the importance that ELR focuses on aggregate data, while lawyers often deal with individual cases. They also show that research on estimator variables when dealing with eyewitnesses (like stress, lighting levels and the presence of a weapon) and system variables ('the features of the identification that are under the control of the legal system') can shed light on individual cases, although the primary goal is to find evidence 'about factors that affect the ability of eyewitnesses *generally* to make accurate identifications'.
18. In medical studies Lasagna's law is well-known, basically saying that researchers have a tendency to overestimate the pool of available patients who meet the inclusion criteria and would be willing to enroll into the study.
19. A well-known statement by Guba (1987: 31) was that 'the one (paradigm) precludes the

- other just as surely as belief in a round world precludes believing in a flat one'. Howe (1988: 10) labeled these thinkers as 'advocates of the "incompatibility thesis" – the belief that the compatibility between quantitative and qualitative methods is merely apparent'.
20. See Greiffenhagen et al (2011) for a methodographic study on differences between the two 'paradigms' in practice and Onwuegbuzie et al (2012) for innovations in qualitative research methods that bridge the gap between 'Quant' and 'Qual'.
 21. Those that are not referred to in footnotes can be found in Cane and Kritzer (2010).
 22. See Eshuis et al (2012) and Eshuis and Leeuw (2014).
 23. It started in the world of medicine with the Cochrane Collaboration/Library, now covering hundreds of thousands (experimental and other types) of studies in medicine and public health. See Chapter 5.
 24. Lawless, Robbenholt and Ulen (2010: 15) say that empirical research is 'an incremental and ongoing process'.
 25. Addressing empirical legal research in a professional way necessitates that attention is paid to *all* aspects of this cycle. Unfortunately, that is not always the case. Brems (2008) studied what 65 (human rights) legal scholars 'from her address book' thought about the importance of methodology for human rights research, how they used methodology, whether or not they were trained in it etc. Unfortunately, the idea of an empirical cycle, including types of research problems, the role of theories and research designs and data analysis, were *not* even mentioned by this author.
 26. The comparison between researchers and bakers is largely incorrect. Many bakers combine the art and craft when making (truffle) cakes, Cote d'Ivoire chocolate cookies and bio-bread, which is not an easy, mechanical activity.
 27. An example of an unprofessional approach to ELR is given by Heise (2002: 821): 'When I speak of empirical legal scholarship I refer only to the subset of empirical legal scholarship that uses statistical techniques and analyses'. Crucial characteristics are missed, as will be clear for the readers of this book (and a few others).