

EVIDENCE ASSESSMENT AND STANDARDS OF PROOF: A MESSY ISSUE*

Giovanni Tuzet

Bocconi University

giovanni.tuzet@unibocconi.it

ABSTRACT: The article distinguishes evidence assessment criteria from standards of proof and addresses three main questions. First: why do some scholars and decision-makers take assessment criteria as standards of proof and vice versa? The answer comes from the fact that some legal systems are more concerned with assessment criteria and others with standards; therefore, jurists educated in different contexts tend to emphasize what they are more familiar with, and to assimilate to it what they are less familiar with. Second: why do systems differ in those respects? Here the answer stems from the historical, institutional and procedural differences that explain why some systems are more concerned with assessment criteria and others with standards of proof. And third, assuming that both criteria and standards are necessary to legal decision-making about facts: how can a system work if it neglects one of these things? Here the article argues that there is a *functional connection* between criteria and standards. The functional connection account is distinguished from a functional equivalence account, and some systems and jurisdictions are referred to in greater detail to support the functional connection claim.

* Earlier versions of this work have been presented in Trento (“GTR 15. Trento Days on Rhetoric”, University of Trento, Italy, June 2015), Washington D.C. (“XXVII World Congress of the International Association for the Philosophy of Law and Social Philosophy”, Georgetown Law Center, July 2015), London (“ICAIL 2017 Workshop”, King’s College, June 2017) and Santiago de Chile (“Seminario 3”, Universidad Alberto Hurtado, June 2017). I thank all the participants in those events who commented on my work; in particular, I wish to thank Daniela Accatino, Ron Allen, Damiano Canale, Flavia Carbonell, Rodrigo Coloma, Christian Dahlman, Marcello di Bello, Paolo Heritier, Mike Pardo and Michele Taruffo. An abridged and earlier version has appeared as Tuzet 2020. Last but not least, I thank the reviewers of this journal for their comments on the submitted draft.

KEYWORDS: assessment of evidence, fact-finding, functional connection, legal proof, standards of proof

SUMMARY: 1. INTRODUCTION.— 2. ANALYTICAL STARTING POINTS.— 3. COMPARATIVE CONSIDERATIONS: 3.1. Common Law Countries. 3.2. Civil Law Countries.— 4. THE FUNCTIONAL CONNECTION ACCOUNT.— 5. CONCLUSION.— 6. BIBLIOGRAPHY

RECOMMENDED CITATION: TUZET G., 2020: «Evidence Assessment And Standards Of Proof: a Messy Issue», in *Quaestio facti*, 2: 87-113. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22480

1. INTRODUCTION

There is a persistent confusion in legal literature. It concerns evidence assessment criteria (or evidence evaluation methods) and standards of proof, which are frequently mixed up and treated as if they were the same thing.

The present work is motivated by the wish to clarify this issue providing an adequate conceptual distinction between assessment criteria and standards of proof. In a nutshell, evidence must be assessed in order to check whether it satisfies a relevant standard of proof; the assessment is operated with some criteria, and both criteria and standards are necessary to fact-finding. This is not as trivial as it may appear because, as I will show, some legal scholars, practitioners and decision-makers seem to miss the difference between criteria and standards.

In addition to this conceptual point, the article addresses three main questions. First: why do some scholars and decision-makers persistently take assessment criteria as standards of proof and vice versa? The answer to this question, simply put, comes from the fact that legal systems are different because, inter alia, some are more concerned with assessment criteria and others with standards; therefore, jurists educated in different contexts tend to emphasize what they are more familiar with, and to assimilate to it what they are less familiar with. Second question: why do systems differ in those respects? Here the answer stems from the development of different perspectives over time: there are historical, institutional and procedural differences that explain why some systems are more concerned with assessment criteria and others with standards of proof. And third, assuming that both criteria and standards are necessary to legal decision-making about facts: how can a system work if it neglects one of these things? The answer to this is the most theoretically ambitious part of the article, since I argue that there is a *functional connection* between criteria and standards; the functional connection account is distinguished from a functional equivalence account, and some systems and jurisdictions are referred to in greater detail to support the functional connection claim.

This article relies on, and is indebted to, a rich literature on evidence assessment and standards of proof. Indeed, there has been a lot of work, over the last years, on

legal standards of proof. Philosophers, legal argumentation theorists, evidence scholars, civil and criminal procedure scholars, among others, have extensively dealt with this issue. Some works make an analytical effort to clarify the idea of a probative standard; others carry out a descriptive study of how standards actually operate; others advance normative claims in the hope of suggesting better standards, or at least better-defined ones; and some do more than one of these things (by defining the issues, describing the existing standards and making normative proposals)¹.

That complex work on standards has been mainly, but not exclusively, carried out in the common law world. In the civil law world, there is less attention to that issue, but a greater concern for the criteria of evidence assessment. A very old Continental debate addresses the question whether it is better to have evidence assessed according to predetermined and fixed criteria (the idea of a «legal proof» system), or according to more relaxed ones accommodated by fact-finders themselves (the idea of a «free assessment» of evidence)². A way of stressing this difference consists in saying that fixed criteria are objective and relaxed ones are subjective, in that the former constrain the fact-finders in a way that the latter do not. However, it is also possible to understand the latter as objective: their objectivity would not consist in the fact that they are legally mandated in virtue of some authoritative source, but in the fact that they put some epistemological constraints on legal decision-making³.

Now, looking at these issues and trying to conceptualize them, it becomes clear that *an evidence assessment criterion is one thing, and a proof standard is quite another*. And both are necessary to legal fact-finding, because evidence assessment is not sufficient to make a decision and because, in principle, a proof standard can be met with different criteria. Consider the «beyond a reasonable doubt» standard: in principle it can be met both with fixed and relaxed criteria, provided that in the first scenario the competent legal authority determines *ex ante* the fixed value of the relevant types of evidentiary items, whereas in the second scenario fact-finders make appeal to non-fixed criteria (being epistemic, practical, or other) to assess the evidence at disposal and judge whether it satisfies the relevant standard.

Still, as previously mentioned, both in the common law and in the civil law world there is some tendency to overlap those issues and mix them up. For instance, the French *intime conviction* is often taken as a subjective criterion of assessment, while others take it as a subjective standard of proof (more on this below). A dispute that occurred some years ago in a top journal still maintains its significance in this respect⁴.

¹ See, among others, REDMAYNE, 1999, LILLQUIST, 2002, PICINALI, 2013. See also VÁZQUEZ, 2013.

² On the criminal side of the debate see FERRAJOLI, 1989: 112 ff. See TARUFFO, 1992: 361 ff on civil aspects too.

³ See TARUFFO, 1992: 370 ff and FERRER 2007: 147 ff.

⁴ See CLERMONT and SHERWIN, 2002 and TARUFFO, 2003a.

The structure of the present article is as follows: § 2 makes some analytical starting points to clear the discussion ground; § 3 performs some comparative work considering different legal systems and the ways in which they deal with the issues discussed here; § 4 tries to convey a sharper analysis and argues that criteria and standards have a functional connection; § 5 concludes by recalling the questions addressed and summing up the answers provided.

2. ANALYTICAL STARTING POINTS

Evidence presented at trial enables fact-finders to reach accurate verdicts on litigated facts. Evidence is necessary to support the factual claims made by the parties and the findings of fact made by the decision-makers. But evidence per se does not yield verdicts. To this purpose there are at least five requirements on juridical evidence: 1) evidence must be admissible according to the rules of the relevant legal system; 2) it must be presented to fact-finders through some ostensive act; 3) it must be «inferentialized» by the parties and the fact-finders, since evidence doesn't speak for itself and the participants in a dispute have to construct evidentiary arguments based on the items presented; 4) it must be assessed to determine its probative value—or, better, the evidentiary inferences and arguments at stake must be assessed to determine the evidentiary support, or warrant, or justification provided by the premises to the conclusions; 5) fact-finders need to consider whether the evidence meets the relevant standard of proof, or whether the relevant burden of proof has been satisfied⁵.

The points that concern us here are the fourth and the fifth. It is one thing to assess, or to evaluate, the evidence presented. And quite another to find whether it satisfies the relevant standard of proof. The two operations are conceptually different and the latter requires the former as a necessary condition⁶. No one can determine whether the evidence admitted and presented meets the relevant standard, unless the evidence is assessed. To be more precise, once the evidence has been «inferentialized», namely translated into evidentiary arguments, the fact-finders have to assess such arguments in order to see whether the relevant standard is met or not. On the other hand, assessment per se doesn't determine a decision either, because a definite outcome can be justified only if a standard of proof is considered.

So, both assessment criteria and standards of proof are necessary conditions of legal decision-making about facts, and none of them is a sufficient condition of it.

⁵ Many of the points discussed in this article under the heading of «standards of proof» can be also discussed under the heading of «burdens of proof» (see ALLEN, 2014). To make things simpler I won't use the latter terminology and the corresponding conceptual apparatus. (However, to my understanding, the burdens terminology is more focused on procedural and argumentative features, while the standards terminology emphasizes probative and epistemic issues.)

⁶ Notice these different but related points: assessment is necessary for the judgment on standard satisfaction, and both are necessary for a justified decision-making.

This appears quite obvious from a conceptual and philosophical point of view. But, surprisingly enough, some jurists and courts seem to miss the point. Some scholarly literature overlooks the issue, and some decision-makers seem to mix things up, which is more concerning. One significant example is provided by the European Court of Human Rights (ECtHR). In the 2010 case *Gäfgen v. Germany*, for instance, the Court says that «in assessing the evidence on which to base a decision [...] the Court adopts the standard of proof “beyond reasonable doubt”» (§ 92). Apparently in this⁷ and similar decisions⁸ the standard of proof is taken to be an evidence assessment criterion, instead of being a threshold for decision which comes into play once the evidence has been assessed. The same confusion is not infrequent in scholarly work, when evidence assessment criteria are equated with standards of proof or vice versa⁹. One scholar, for instance, asks this question: «Is the standard of proof in the Continental systems, sometimes called *intime conviction* according to its French variant or *freie Überzeugung* in German, really lower than the standard “beyond reasonable doubt” of which common lawyers tend to be so proud?»¹⁰. This question reveals the confusion I am addressing.

Now, why is that? Why do educated scholars and skilled decision-makers fall into this confusion about evidence and proof? Why do they fail to see the difference between assessment criteria and standards of proof? Before advancing my hypothesis on this somewhat surprising phenomenon, let us take a closer look at the rules of some contemporary systems.

⁷ The interest of *Gäfgen* comes also from the fact that it deals with a puzzling «fruit of the poisonous tree» case: conclusive evidence of guilt was obtained as a result of a confession that was allegedly coerced in that the applicant had been threatened by the police for the purpose of finding a boy he had kidnapped. On this case see JACKSON and SUMMERS, 2012: 192-193.

⁸ The understanding of «beyond reasonable doubt» as an assessment criterion appears as a recurrent theme in the decisions of the ECtHR; see also, e.g., *Ireland v. UK* (1978), § 161; *Labita v. Italy* (2000), § 121; *Fedorov v. Russia* (2011), § 57; *Rizvanov v. Azerbaijan* (2012), § 45; *Najafli v. Azerbaijan* (2012), § 36. On the other hand, in these decisions the Court usually adds that such proof may follow from sufficiently «strong, clear and concordant inferences» or «similar un rebutted presumptions of fact», which in fact amount to ways of assessing the evidence (epistemically as to «strong, clear and concordant inferences» and legally as to presumptions). Instead, the Inter-American Court of Human Rights does not apply an explicit standard but according to some commentators (e.g. BOVINO, 2005 and PAUL, 2012) it generally uses the preponderance of the evidence (civil) standard, for its purpose is not the punishment of human rights violations but the protection of victims and the «reparation of damages».

⁹ On the common law side, see, e.g., CLERMONT and SHERWIN, 2002 (treating *intime conviction* as a standard); on the civil law side, see, e.g., IACOVIELLO, 2006 and CAPRIOLI, 2009 (treating the «beyond a reasonable doubt» standard as an evidence evaluation method because of the judicial obligation to deliver written reasoned opinions). Of course, there are also scholars who see well the difference (especially in the Spanish-speaking world: Daniela Accatino, Rodrigo Coloma, Jordi Ferrer, Marina Gascón in particular).

¹⁰ BOHLANDER, 2012: 6.

3. COMPARATIVE CONSIDERATIONS

I sketch in this section an overview of the rules of some actual legal systems about evidence evaluation and proof standards. Some systems have rules that clearly consist of standards of proof, whereas others have rules that clearly consist of evidence assessment criteria, while there are also rules which are difficult to categorize one way or the other. Given the theoretical purpose and the conceptual approach of the present article, I will overlook many positive details and make abstraction from many systemic complexities. What is important for the purposes of this work is the grasping of the ways in which different systems handle the messy issue we are dealing with.

I shall start by making some remarks on the common law world and then move to some civil law countries.

3.1. Common Law Countries

In the common law world there are three basic standards of proof. Two of them apply to civil cases, one to criminal cases. In the following I refer especially to the systems of the US and the UK (England and Wales in particular).

In civil cases the usual standard is the *preponderance of the evidence* (as they call it in the US), or the *balance of probabilities* (as they call it in the UK)¹¹. This means that, according to the received view, the claim of the burdened party (usually the plaintiff) must be more probable than not in light of the evidence presented. If, given the evidence, the factual claim of the party with the burden of proof is considered to be more probable than not, that party is entitled to a verdict in its favor. If it is not considered to be so, the party has no such entitlement.

But in some civil cases the US system adopts a second standard, the *clear and convincing evidence* standard. It is used in «serious cases», namely those in which what is at stake is more important than in the rest. Examples of such «seriousness» are punitive damages cases—for instance for a fraud charge—and immigration cases where what is at stake is the right to be in a certain country¹². This second standard is more demanding than the first. In order to have a verdict in its favor, the burdened

¹¹ By the way, it is possible to speculate about some differences: one of these might be that the UK formula explicitly suggests a probabilistic account of evidence, which the US formula does not; similarly for the «more probable than not» version of the standard, which is even more committed to some probability account, namely to an account of probability that assumes the complementation rule according to which, if the probability of H is p, then the probability of not-H is $1 - p$ (which is not necessarily the case with the «balance of probabilities» version of the standard, and even less so with the «preponderance of the evidence» version if probabilities are understood epistemically rather than mathematically; see COHEN, 1977 and HAACK, 2014: 47 ff).

¹² See REDMAYNE, 1999: 189 (suggesting the adoption of this standard in the UK to face such serious cases).

party must make a greater evidentiary effort, presenting evidence which is stronger, or which has a higher probative value, than that which suffices in the rest of cases.

In criminal matters the well-known standard of guilt, especially entrenched in the US¹³, is the proof *beyond a reasonable doubt* (BARD). This standard requires an amount of evidence that only leaves room for unreasonable doubts about the defendant's guilt. So, if the evidence presented makes it reasonable to believe that the defendant is guilty and makes it unreasonable to doubt it, then the prosecution is entitled to a verdict in its favor and the triers of facts are committed to decide against the defendant. Of course this criminal standard is even more demanding than the clear and convincing evidence standard used in serious civil cases: the evidence presented has to be even stronger than this, leaving room for unreasonable doubts only, which is to say doubts that lack specific epistemic reasons, as spectacularly exemplified by merely skeptical doubts (such as Cartesian doubts and «brains in a vat» doubts¹⁴) or motivated by essentially philosophical concerns cast in probabilistic terms (such as the idea that we cannot be 100% sure of anything¹⁵).

There have been some attempts to translate those qualitative standards into quantitative thresholds. This is done with some probability theory and some assignment of a numerical probability value to each of the standards. It is not a difficult task if we consider the preponderance of evidence, which is easily translated into something such as $>.5$, while the clear and convincing evidence is usually considered to be around $.7$ (or $.75$) and the proof beyond a reasonable doubt is usually located around $.9$ (or $.95$)¹⁶.

The rationale for having a higher standard in serious civil cases, compared to non-serious ones, is the protection of some more important value which is at stake in them. And the rationale for having an even higher standard in criminal cases, compared to civil ones, is the protection of a fundamental value such as individual liberty (and even life in those systems where capital punishment is still in place). This is fairly obvious and there are quite sophisticated pieces of literature that work out the details of this plain remark¹⁷. What is less obvious is the capacity of those standards to do the job they are designed for, due to the lack of empirical information about

¹³ It is so by virtue of some Supreme Court's decisions such as *Miles v. United States*, 103 U.S. 304 (1880) and, in particular, *In re Winship*. 397 U.S. 358 (1970).

¹⁴ PUTNAM, 1995: 17 and PUTNAM, 1981: 1 ff.

¹⁵ A traditional phrase used in various contexts to qualify the certainty required for a guilt verdict is «moral certainty», which falls short of «absolute certainty» and sounds like a standard rather than a criterion.

¹⁶ Of course, the $.9$ value is worrisome for criminal defendants, but the higher the standard is set (to avoid false convictions) the higher is the expected rate of false acquittals. Consider also that, when a 5/6 jury majority is required for conviction (as it was according to the French 1791 system), the corresponding probability is even lower, namely $.83$ (see PADOA-SCHIOPPA, 1987: 92, 142). I cannot expand on these topics here.

¹⁷ See especially BELL, 1987. Cf. STEIN, 2005: 118 ff and LAUDAN, 2006: 63 ff.

particular cases or to conceptual confusion about some aspects of the standards. There is in fact an increasing amount of critical literature (on the criminal standard in particular¹⁸) motivated by the idea that such standards are not really helpful in legal decision-making, insofar as they do not provide genuine epistemic reasons to decide a case one way or the other. This is not the topic of the present article, though. What I want to focus upon is the ultimate nature, so to say, of the standard of proof rules: they are *decision rules*.

The standard of proof rules set the qualitative or quantitative thresholds that must be reached to have a decision in favor of the burdened party. This party bears the risk of there being an amount of evidence which is insufficient with respect to the standard. If the evidence is insufficient, the burdened party has no entitlement to a decision in its favor. If it is sufficient, the party has such an entitlement. So standard of proof rules are in the end decision rules. They justify decisions in favor of one party or the other, based on the factual claims and the arguments given the evidence.

Now, all of that presupposes an assessment of the evidence. An evaluation of the evidence which was admitted, presented and discussed is necessary to the judgment about standard satisfaction. Standards of proof, pace some scholars and judges, are not criteria of evidence assessment. Instead, they presuppose and require some criteria that determine the probative force, or probative value, of the evidence in play, so as to see whether it meets the relevant threshold.

To make a simple exemplification, it is not enough to assign the plaintiff's claim a .65 probability of being correct: in order to make a decision we need to know whether the standard is the simple preponderance of the evidence or the clear and convincing evidence standard. On the other hand, the standard itself does not tell you if the evidence presented satisfies it: you need to assess the evidence, either quantitatively or qualitatively. To take a qualitative example, the well-known decision in O. J. Simpson's criminal case can be understood like this: the assessment led to the «he did it» hypothesis as the best explanation of the evidence but it also led to an acquittal given that, according to the jury, the evidence didn't match the BARD standard.

If this reconstruction is correct, where are the evidence assessment criteria in the common law world? Basically, they are in the scholarly disputes about evidentiary value, or probative force, or the weight of (combined) evidence¹⁹. They are not cod-

¹⁸ LAUDAN, 2006: 29 ff. Note that the UK has been recently abandoning BARD, or at least its wording (out of troubles in making it understood): now, following the Judicial College (formerly the Judicial Studies Board), UK fact-finders are to convict only if they are *sure* of guilt (see R. v. Majid [2009] EWCA Crim 2563; see also ROBERTS and ZUCKERMAN 2010: 253 ff). However, on BARD's interesting origins in England and Continental Europe, see WHITMAN, 2008 (it was supposed to protect the juror's religious conscience facing the risk of wrongful convictions, it was not originally designed to protect the defendant).

¹⁹ See, among others, ANDERSON, SCHUM & TWINING, 2005: 224 ff; BARZUN, 2008; HAACK, 2014: 208 ff.

ified, as far as I know, in statutory rules nor in any other legislative materials. Truly, there are some documents that give general guidance. Consider for instance the Crown Court Compendium (Part I, on «Jury and Trial Management and Summing Up») by the English Judicial College, where assessment criteria for the jury are given with reference to inferences from expert evidence, bad character evidence, hearsay, and the defendant's silence at trial, to mention some of the issues. For the rest, the criteria are advanced in scholarly controversies and reflected in jury instructions and judicial opinions insofar as judges and courts take a position on such controversies. Consider the literature about probability theories and their application to legal matters and cases; consider the controversy between the supporters of some subjective (Bayesian) conception of probability and the supporters of objective probabilities, or the dispute between Pascalian and Baconian probabilities, or the passionate controversy between the supporters of the relative plausibility account and the supporters of the mathematical probability account, that is, between an ordinal assessment of factual claims and a cardinal one²⁰; consider in addition the debate between atomism and holism in evidence assessment²¹; and consider also the ways in which judges explicate to juries, with the so-called «jury instructions», how to evaluate whether a standard is met. All of this is about the criteria that fact-finders need to use to check whether the evidence presented in a case satisfies the relevant standard of proof.

3.2. Civil Law Countries

Civil law countries usually have in their respective codes some basic rules about evidence and proof. Most of the time these rules set evidence assessment criteria instead of standards of proof. And nowadays these criteria are rather relaxed, leaving room to judicial discretion. This has a definite historical explanation: in many Continental countries the codification effort pointed at getting rid of, inter alia, the prior complexities in the assessment of evidence. Those complexities characterized Roman Canon law and amounted to the so-called system of «legal proof», which lived from the end of the XIII to the end of the XVIII century up to the French Revolution, and which dominated Europe in the XVI and XVII century in particular²². In that system every sensible kind of evidence was assigned a predetermined and fixed probative value, on the basis of some legal authority or doctrinal opinion; then, once some evidence was presented at trial, the judge's task was basically that of calculating the value of each factual claim given the evidence at disposal. In other words, once the *types* of evidence had received a fixed value, the judges were to consider the *tokens*

²⁰ See, e.g., TRIBE, 1971, COHEN, 1977, TILLERS and GREEN, 1988, ALLEN, 1991, GOLDMAN, 2002, REDMAYNE, 2003, ALLEN and PARDO, 2007, KAPLOW, 2012, ALLEN and STEIN, 2013, VERHEIJ et al., 2016, ALLEN and PARDO, 2019.

²¹ See DAMAŠKA, 1997a: 34-37, HAACK, 2014: 235-238 and SCHWEIZER, 2014.

²² See among others DAMAŠKA, 1986: 29 ff, PADOA-SCHIOPPA, 2003: 280-292, JACKSON and SUMMERS, 2012: 57 ff.

in the case at hand and were to calculate the outcome accordingly, with no or little discretion on their part²³. The best-known example of that is the traditional two-witness rule according to which two independent witnesses testifying the same amount-ed to «full proof» of the fact so testified, leaving no room to the judicial appreciation of the witnesses' credibility, of the circumstances of testimony, etc. And at the same time, according to that rule, a single testimony was not sufficient to prove a claim, however credible and accurate the testimony might have been. Then, to cut a long story short, Enlightenment legal thinkers and French Revolutionary authorities designed a system that, on the contrary, trusted fact-finders conferring some discretion upon them but, at the same time, stating guidelines of evidence assessment²⁴. Eventually the spirit of the Revolution spread throughout Europe and civil law countries adopted the new framework, which is still in place.

Let me start with Italy. The actual code of civil procedure states that fact-finders must evaluate the evidence according to their «prudent assessment» (*prudente apprezzamento*, Art. 116), unless a criterion of «legal proof» (*prova legale*) is mandated. There are in fact some provisions attributing legal value to some kinds of evidence such as official documents (Art. 2700). But these are the exceptions to the general rule of «prudent assessment». This rule provides a general criterion of evidence assessment. Then the question is: what is the standard of proof in the Italian system of civil procedure? It is hard to say, because there is no explicit rule neither in the code nor in statutory law. In the last years—against the view that the standard is something like «moral certainty» or even full conviction—some courts have felt the need to specify that the standard of proof in matters like medical malpractice is the «more probable than not» idea²⁵. So, according to these judges, it is basically the standard used in common law countries. On the other hand, it is unclear whether the standard is conceived to apply only to such tort matters as medical malpractice or more generally to any matter of civil law and procedure.

The situation is different if we take into consideration the Italian criminal procedure code: fact-finders shall convict only if the BARD standard is met (Art. 533, as modified in 2006)²⁶. The standard was introduced by Parliament some years ago

²³ But see DAMAŠKA, 2003: 129: «It is true that these lawyers developed a maze of rules about the quantity and quality of evidence needed for fact determinations. But it is a mistake to believe that these rules turned adjudicators into automatons, who made factual determination on the authority of rules, independently of their beliefs». On the other hand, DAMAŠKA, 1986: 55 points out that «even today the Continental “free evaluation of evidence” is not really free: as befits the hierarchical process, trial judges are required to justify their findings of fact, and the cogency of their reasoning is scrutinized by appellate courts». Cf. DAMAŠKA, 2018.

²⁴ A key event was the adversary reform and the introduction of the criminal jury in France by the law of September 16-29, 1791 (statute of September 16, decree of September 29).

²⁵ See, e.g., Cass. sez. III civile, decision n. 10285/2009, n. 10741/2009, n. 15991/2011, n. 23933/2013, n. 18392/2017.

²⁶ A nice aspect of the Italian formula is that it reads *oltre ogni ragionevole dubbio* («beyond all reasonable doubt»). There isn't any logical difference with the «beyond a reasonable doubt» formula (for

with the purpose of making criminal convictions harder. And it was supported by the belief that Americans better protect criminal defendants. Some commentators objected, saying that the idea was already implicit in the system, and that therefore that legislative intervention was superfluous²⁷. But this is not our concern here. The interesting question is, given the now explicit standard, whether the code mandates any assessment criterion: in fact, as a general matter, there is just one laconic provision according to which the judge assesses the evidence giving an account of the criteria adopted and the results obtained (Art. 192 c. I). Literally construed, this provision makes any assessment criterion acceptable, provided that the judge specifies the method used and the results arrived at. Anything goes, apparently. But this would be an inappropriate construal of the text; for the system purports to be rational and to have evidence rationally assessed²⁸. To take a somewhat extreme example, the judge could not say something like this: «I have consulted a clairvoyant and reached the conclusion that the defendant is guilty». Consulting a clairvoyant would be the method used, and the defendant's guilt would be the result arrived at. So, literally speaking, the judge would comply with the rule that requires making method and results explicit. But that would run counter to the goal of rationality. Additionally, there are more definite rules concerning things like circumstantial evidence (that must be «serious, precise and consistent» to prove a fact, according to Art. 192, c. II) and the declarations of the co-defendant (that must be assessed considering the evidence that corroborates their reliability, according to Art. 192, c. III). So one can point out that Italy has a criminal procedure system with an explicit standard of proof but, as a general matter, no explicit assessment criterion, whereas the civil procedure system has it the opposite way, for it has an explicit assessment criterion and no explicit standard of proof.

Let's consider France now. According to the French civil system there is «free proof» of legal facts (see Arts. 9-10, 179, 198, 213 of the code of civil procedure)²⁹ and «legal proof» of (certain) legal acts (see, e.g., Arts. 1359, 1371 of the civil code). No explicit standard is given, and the situation is similar to the Italian civil one.

The most interesting thing of the French system is the highly controversial notion of *intime conviction* used in criminal trials. According to the criminal procedure code, judges must decide on the basis of the evidence according to their *intime conviction* (Art. 427); the same holds for mixed courts, made of judges and jurors who

here «a reasonable doubt» means «any of the doubts that are reasonable»); but the rhetorical emphasis of the «all» formula appears to be heavier. Similarly for the Chilean formula (see below in the text) «más allá de toda duda razonable».

²⁷ See the reconstruction of the debate in DELLA TORRE, 2014: 9-11.

²⁸ The fundamental right to have evidence examined and the other «fair trial» (*giusto processo*) requirements militate for this conclusion. See Arts. 24 and 111 of the Italian Constitution, Art. 6 of the European Convention on Human Rights, and Art. 47 of the EU Charter of Fundamental Rights.

²⁹ See also Art. 1381 of the French civil code on the probative value of testimony, to be determined by the judge.

are not bound by legislative rules of assessment and must ask themselves «in silence and meditation» (*dans le silence et le recueillement*) if, given the evidence, they have an *intime conviction* about the defendant's guilt (Art. 353; see also Art. 304). There has been a lot of talk about this idea. Several scholars have criticized it assuming it is nothing but an appeal to the fact-finder's psychology³⁰. If it were so, it would not provide any epistemic ground for the verdict and it would open the door to any form of bias, suggestion and emotive decision-making. Now, frankly, this is an uncharitable reading of the rule. The relevant texts underscore the need to fix that state of mind *on the basis of the evidence*; it is not evidence-free, it does not boil down to the purely subjective and psychological feeling about the defendant's guilt. It is true that the appeal to «silence and meditation» strikes an Anglo-American scholar as utterly different from the idea of heated disputes between jurors, with arguments and counterarguments being debated until a decision is made. It strikes even a civilian like me in that it appears as the «monk way» to the discovery of truth, or, philosophically speaking, the «Cartesian way» to legal fact-finding (as is well known the dimension of solitary meditation is a peculiar feature of Descartes' philosophy). But *intime conviction* should not be taken as the legitimation of any kind of subjectively formed or emotionally driven decision³¹. Originally it was adopted with the idea of transplanting into the French system the decision-making way of English juries³². Given the Enlightenment trust in the judgment of lay people (and a lesser trust in the opinion of professionals like judges) the French system adopted the jury method for criminal trials and set *intime conviction* as the check on the defendant's guilt. Eventually the idea of *intime conviction* was extended to mixed courts and to professional fact-finders such as judges.

So, is *intime conviction* a standard of proof? Is it rather an evidence assessment criterion? Is it both? Anglo-American commentators tend to think it is a standard³³. Ironically, French jurists and civil lawyers tend to think it is not a threshold, being instead a way to assess the evidence and make a decision³⁴. And even reputed scholars and evidence theorists of the civil law world disagree on the understanding of it; for instance, Jordi Ferrer thinks it is an «entirely subjective» standard of proof,

³⁰ See, e.g., TARUFFO, 2003b: 81-82, LAUDAN, 2005: 98-99, FERRER, 2007: 144-145, BAYÓN, 2008: 17-18. An echo of this criticism can be found in some judicial opinions in different countries: see, e.g., Tribunale di L'Aquila, sez. penale, decision n. 380/2012, § 4.1 (which is a decision on the criminal charges relating to the L'Aquila earthquake of 2009).

³¹ HANS and GERMAIN, 2011: 755: «the decision based on *intime conviction* is not best viewed as the expression of a feeling, but rather as a considered opinion based on the charges, evidence, and defenses presented by the parties».

³² See PADOA-SCHIOPPA, 1987 and 1994.

³³ See, e.g., CLERMONT and SHERWIN, 2002: 256. Cf. HANS and GERMAIN, 2011: 754.

³⁴ See, e.g., BREDIN, 1996: 23, DELMAS-MARTY, 1996: 59-60, BOULOC, 2004: 54, CARTIER, 2004: 70, BERGEAUD, 2010: 160; see also TARUFFO, 2003a: 666, CASTILLO DE LA TORRE, 2009: 522 and UBERTIS, 2013: 332.

whereas Daniel González Lagier says it is a «system of evidence evaluation»³⁵. To my understanding the reading of it as a standard is in fact supported by the presence, in the same code, of a provision according to which judges are to assess freely every evidentiary item (Art. 428). This is an evaluation criterion and if *intime conviction* were a criterion itself the system would be either redundant or misleading.

Let us turn to Spain now. According to the civil procedure code, evidence must be assessed according to the «rules of sound criticism» (*reglas de la sana crítica*, e.g., in Art. 348 on expert testimony, and in Art. 376 on witnesses). The situation is similar to that of other civil law countries in that «legal proof» is the exception, not the rule. Here the rule is assessment according to «sound criticism». Some commentators appreciate this criterion insofar as it is more epistemologically inclined than the mere idea of a «free assessment»³⁶. Many Spanish-speaking countries in Latin America adopt the same criterion. Such an epistemological flavor is missing instead from the Spanish code of criminal procedure, according to which judges decide assessing the evidence with their «conscience» (Art. 741.1). The appeal to conscience seems to be in line with the French appeal to *intime conviction*. Other Spanish-speaking countries follow the same path, except Chile where both «sound criticism» and «beyond a reasonable doubt» are used in criminal cases (see Art. 369-bis of the criminal code and Art. 340 of the criminal procedure code)³⁷. The case of Chile is interesting because it provides a confirmation of the conceptual intuition that assessment criteria and standards of proof are different things. Chile adopts «sound criticism» as the evidence assessment criterion to be used in criminal cases, and «beyond a reasonable doubt» as the standard of proof to be used there. This double discipline is not redundant, for criteria and standards are different³⁸.

Finally, a few words on Germany. According to the civil procedure code, the court is to assess the evidence freely (*Freie Beweiswürdigung*): it is supposed to decide «at its discretion and conviction» and taking into account «the entire content of the hearings and the results obtained by evidence being taken» (Sec. 286.1); the court «shall be bound to statutory rules of evidence only in the cases designated in the present Code» (Sec. 286.2). The emphasis on judicial discretion confirms the trend of civil law systems in civil matters³⁹. Coming to the German code of criminal

³⁵ FERRER, 2007: 145 and GONZÁLEZ LAGIER, 2013: 52.

³⁶ COLOMA and AGÜERO, 2014; LUCH, 2012: 194-196; TARUFFO, 2003b: 82.

³⁷ See also Art. 297 of the criminal procedure code, that establishes the free assessment of evidence in compliance with logic, experience and science. This provision does not mention «sound criticism» but is usually read as a reformulation of this traditional idea; the previous code of criminal procedure (substituted by the present one in 2000) mandated an evaluation of the evidence according to the «rules of sound criticism» (Art. 562).

³⁸ See ACCATINO, 2011 and COLOMA, 2012. See GONZÁLEZ, 2006 on civil aspects too.

³⁹ Just before the German reunification a scholarly dispute occurred on the alleged «German advantage in civil procedure»: see LANGBEIN, 1985, ALLEN et al., 1988, LANGBEIN, 1988 and ALLEN, 1988. The dispute is still interesting for a cost-benefit analysis of party-dominated and judge-dominated systems of procedure and fact-gathering in particular. See also SCHWEIZER, 2016 on the civil

procedure, the court shall decide «on the result of the evidence according to its free conviction gained from the hearing as a whole» (Sec. 261). Again, the emphasis is on free assessment and free conviction (*Freie Überzeugung*)⁴⁰. But still on the basis of the evidence and the hearing. So, both in the civil and criminal context German law emphasizes the role of judicial discretion in evidence assessment. No standard of proof is made explicit in positive law, unless one reads «free conviction» as a standard, which is puzzling as much as the idea of having a «free standard». But the attentive reader would have noted that those provisions make reference to decision, not only to evidence assessment. So, such evidence assessment criteria are supposed to play also the role of decision rules.

To summarize, contemporary civil law countries are similar to one another in that, with respect to civil matters, «legal proof» is the exception and «free assessment» of evidence is the rule. According to some scholars there are degrees of such freedom, for the French way is somewhat extreme in giving discretion to fact-finders, while the Spanish way puts more epistemological constraints on their judgment, and the Italian way of «prudent assessment» appears to be in the middle. Michele Taruffo has claimed that «free assessment» is not to be taken as freedom from the rules of logic, the requirements of rationality, and the constraints that science and epistemology put to the triers of fact⁴¹. It is simply freedom from predetermined and fixed rules of evidential weight; freedom from the old system of «legal proof» (with the few exceptions that still remain). It is the idea of having evidence assessed case by case by sensible fact-finders, by working minds that are able to appreciate the particularities of concrete cases and to make responsible factual determinations. Contemporary civil law countries display a lower degree of similarity with respect to criminal matters, since some of them (like Chile and Italy) have explicitly adopted the BARD standard of proof, while others just stick to the Continental idea of a «free» assessment of evidence⁴².

standard of proof in Switzerland and Germany (claiming that it is lower than what doctrine and case law proclaim, namely «full conviction»).

⁴⁰ See, e.g., BOHLANDER, 2012: 8 ff, 32.

⁴¹ TARUFFO, 2009: 160 ff, 219. See also GASCÓN, 2010: 32.

⁴² There are of course further interesting things I cannot examine here: for instance, Sweden has a unique code of procedure (concerning both civil and criminal matters) which mandates the free evaluation of evidence (Chap. 35, Sec. 1) and is quite generous as to evidence admission (see DAHLMAN, 2015 on the use of prior convictions as evidence of guilt); in the Netherlands Sec. 338 of the criminal procedure code mandates free evaluation through the hearing, but Sec. 342.2 states that the court may not find there is proof that the defendant committed the offence as charged in the indictment exclusively on the basis of the statement of one witness (see MACKOR, 2016 on the ways in which the Dutch Supreme Court and some scholars have tried to complement this *unus testis* rule); in Slovenia no explicit standard is mandated, but in criminal matters the Supreme Court has introduced BARD with the holist assessment criterion of interconnected circumstantial evidence (see NOVAK, 2016); in Poland Sec. 233.1 of the code of civil proceedings gives discretion in the evaluation of evidence but requires it to be holistic (I've been told so by M. Araszkiwicz, whom I thank).

The following tables summarize the above findings (the considered countries are in the columns⁴³; in the rows «AC» stands for «assessment criterion» and «SP» for «standard of proof»; in the cells, «?» stands for a missing or unclear aspect, «POTE» stands for the «preponderance of the evidence», «BARD» stands as usual for «beyond a reasonable doubt», and «IC» stands for *intime conviction* or a similar mental state):

Tab. 1. Civil Cases

| | US | UK | I | F | S | C | G |
|----|------|------|---------|------|-----------------|-----------------|------|
| AC | ? | ? | Prudent | Free | Sound Criticism | Sound Criticism | Free |
| SP | POTE | POTE | ? | ? | ? | ? | ? |

Tab. 2. Criminal Cases

| | US | UK | I | F | S | C | G |
|----|------|-----------|------|-----|-------------|-----------------|------|
| AC | ? | ? | Free | IC? | Con-science | Sound Criticism | Free |
| SP | BARD | Certainty | BARD | IC? | ? | BARD | IC? |

A few systems among those considered in these tables are fully explicit (just Italy and Chile on criminal procedure); the majority are half-explicit, for some of them (common law ones) only have explicit standards of proof, and the others (civil law ones) only have explicit assessment criteria.

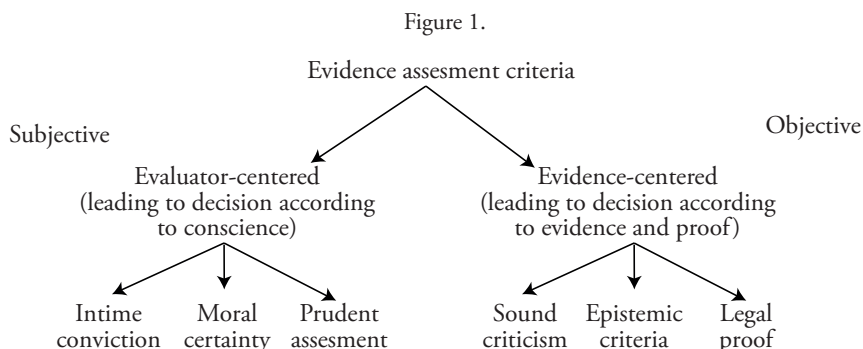
Overall it is not an equilibrium situation, because the general discretion of evidence assessment and the lack of explicit standards in some contexts generate not only additional litigation and controversy but also confusion as to what is missing in, and prescribed by, a given legal system (like the civil procedure one in Italy, for instance).

As to assessment criteria, in a comparative scheme they can be ordered according to the degree of freedom they confer upon the fact-finders, stipulating that a criterion is as «objective» as it decreases such freedom and is as «subjective» as it increases it. So it is not a categorical distinction, it is a matter of degree. Maximal objectivity minimizes judicial discretion, and subjective criteria maximize it⁴⁴. On one side we can locate the fixed criteria of legal proof; on the opposite side we can locate *intime conviction* (if it's read as a subjective criterion); next to legal proof we can locate epis-

⁴³ To avoid any misunderstanding: «US» stands for the United States, «UK» for the United Kingdom, «I» for Italy, «F» for France, «S» for Spain, «C» for Chile, and «G» for Germany.

⁴⁴ Even though it may happen that in practice there are «subjective patterns» or intersubjective ones.

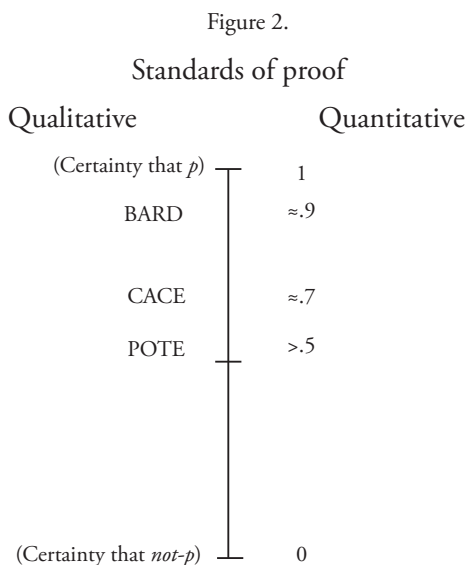
temic criteria, and moral certainty next to *intime conviction*; in the middle, sound criticism is closer to objectivity and prudent assessment is closer to subjectivity (see Figure 1). Probabilistic accounts can be located too in the scheme, according to the probability theory they endorse (subjective v. objective, Pascalian v. Baconian, etc.).



In other words, objective criteria are evidence-centered since they focus on the evidence at disposal, while subjective criteria are evaluator-centered since they focus on the states of mind and attitudes of the fact-finder (on prudence, conscience, conviction, etc.). This is also in tune with the subject-matter of the old controversy over decision according to conscience (*secundum conscientiam* in Latin) or according instead to evidence and proof (*secundum alligata et probata*)⁴⁵.

As to standards of proof, in a different comparative scheme they can be ordered according to the degree of belief on the claim at stake (qualitative reading) or according to the probability value that the claim is attributed (quantitative reading). So, if the claim is p , at one extreme of a vertical segment we have the certainty that p (or probability 1), at the other extreme the certainty that *not-p* (or probability 0), and in the middle the different degrees of belief about p and the three basic standards, where «BARD» stands for «beyond a reasonable doubt», «CACE» for «clear and convincing evidence» and «POTE» for «preponderance of the evidence» (see Figure 2). (Note that this assumes the conventional, or Pascalian, view on probability; a Baconian view would need a different scheme.)

⁴⁵ For instance, in a medieval society it was not unlikely that the same person acted as a confessor and as an ecclesiastical judge; then a problem was posed by the information the judge received as a confessor, especially by a crime confession: were such judges permitted to use it in their judicial capacity? So, were the judges to decide according to conscience or to evidence and proof? See PADOA-SCHIOPPA, 2003: 251 ff.



As an additional point, note some practical consequences of the above distinctions. Note that the scope of legal argumentation is significantly reduced when fixed criteria of evidence assessment are mandated, or as far as these criteria are objective. This was pretty obvious in the old system of «legal proof», where any discussion about the witnesses' credibility or the circumstances of testimony was beside the point once the fact-finders were bound by some measure like the two-witness rule. The more discretion fact-finders have, the more prominent is the role of legal argumentation. Parties are supposed to provide persuasive arguments about evidence and fact-finders are supposed to provide arguments that justify their decisions. This is the case in particular when judges have to provide written opinions that (supposedly) justify the decisions they made. *Intime conviction* is frequently ridiculed as epistemologically unsound, but consider how frustrating it is to be bound by fixed criteria that impede any appreciation of the particular circumstances of the case and block any argumentative effort to make sense of the evidentiary items at disposal.

Interestingly enough, inquisitorial systems are historically tied to legal proof criteria⁴⁶; instead, *intime conviction* was introduced in France in the adversarial spirit of English trials, and if criteria are subjective and there is room for argumentation, parties have a fair chance to make their points (the worst scenario for a criminal defendant is an inquisitorial system with subjective criteria of evidence assessment). Remember also that, according to certain surveys⁴⁷, lay triers of fact judging on *inti-*

⁴⁶ See FERRAJOLI, 1989: 112-113.

⁴⁷ HANS and GERMAIN, 2011: 746.

me conviction are more prone to acquittal than professional decision-makers (against repeated cries that subjective criteria are very bad). Hence something like *intime conviction* is not that bad if it is: i) understood in connection with evidence; ii) encapsulated in an adversary system with argumentation coming from both parties; iii) used by lay fact-finders.

Before moving on, let me also stress that fixed criteria are not to be equated with explicit ones: the fixed criteria of the «legal proof» system were of course explicitly mandated by authoritative legal sources, but these can also mandate relaxed or subjective criteria that leave discretion to fact-finders. Whether a criterion is fixed or not is a substantive issue. Whether it is explicit or not is simply a matter of positive law.

4. THE FUNCTIONAL CONNECTION ACCOUNT

A comparative analysis of the foregoing remarks focuses on the relevant systemic differences. Some legal systems have explicit standards of proof (as decision rules) and lack explicit assessment criteria. Some other systems have explicit assessment criteria and lack explicit standards of proof (again, understood as decision rules). In addition, it is hard to say whether the non-explicit components are implicit or undetermined. For instance, it is *prima facie* hard to say whether in the Italian civil procedure system there is an implicit standard of proof or none.

If all of that is correct two further questions arise, after the first we posed above (in § 2) regarding the confusion between criteria and standards in scholarly work and judicial opinions. The second question is: why do systems differ? That is, why do some of them have assessment criteria and lack standards of proof, while others have it the other way round? And the third question is: how can systems work without one of such elements? That is, if assessment criteria and standards of proof are necessary components of legal fact-finding, how can it be the case that a system that lacks one of them can work nevertheless?

The second of the three questions mainly finds its answer in history. The above-mentioned differences between systems (and the relative confusion in scholarly works and judicial opinions) are generated by different historical, institutional and procedural concerns. The present article is not a piece in legal history, so I won't expand on this point. Suffice it to say that the common law world shows a prominent concern for standards of proof and this finds an explanation, I guess, in the features of jury trial. The fact that the triers of fact are lay people, who need instructions from judges but don't provide written reasons for the decisions they make, militates in favor of a simple system where fact-finders have to consider a standard of proof and can be instructed by judges about the meaning of the standard and the ways to evaluate whether the standard is met. This facilitates the task of the jury, and arguably it does so without dictating how jurors are meant to evaluate the evidence.

On the other hand, the civil law world shows a prominent concern for evidence assessment and this finds an explanation in the features of trials with professional fact-finders like judges, who once were supposed to apply the complex system of «legal proof» and eventually, in contrast with the previous binding system, were asked to freely assess the evidence. Notice that these judges are committed to justify their decisions, namely to give ex post written reasons for the final decisions they make both on factual and legal claims; to this purpose it is not sufficient to say something like «We have been persuaded that the burden of proof was satisfied»: it is essential to justify the way the evidence was evaluated, and to understand how a certain degree of conviction was determined, because the justification of the decision depends not only on the satisfaction of the proof standard but also on the correctness of the process. This focus on the assessment process has likely contributed to the neglect of the standards issue in the civil law world.

The puzzling nature of *intime conviction* is less puzzling if we consider it in the perspective of that historical process. The *intime conviction* idea was introduced in France to emulate the English way to decide criminal cases with jurors and eventually it was extended to judges freed from the complexities of the «legal proof» system; so it is Janus-faced in that it reflects different legal cultures and traditions and because it has absorbed, so to say, the different concerns of these cultures and traditions.

This also explains some conceptual confusion about evidence and proof. Civilians often take standards as assessment criteria because they are more familiar with the latter (consider again the ECtHR decision quoted above). And vice versa common lawyers often take assessment criteria as standards because they are more familiar with the latter (think again of their understanding of the French *intime conviction*). But consider that those views are not entirely unjustified as far as there are some functional connections between criteria and standards: some standards incorporate assessment criteria and vice versa. Consider again BARD: it is arguable that the reasonable doubt standard incorporates the idea of a reasonable assessment of the evidence (which is not very informative, but still a form of guidance). Similarly, for *intime conviction*: if it is an assessment criterion, it incorporates the threshold of the fact-finder's conviction as a decision rule.

These last remarks guide us to the answer to the third question asked above. *There is some connection* between criteria and standards, notwithstanding their conceptual difference. What kind of connection? There is a *functional connection* between them which must be distinguished from a sort of functional identity or equivalence⁴⁸. The idea of a *functional equivalence* would consist in the fact that standards and criteria perform the same procedural and systemic function. This idea has some intuitive interest, for it would explain away the puzzle of systems working fairly well without

⁴⁸ For the equivalence idea see LLUCH, 2012: 192 ff. On functional explanation and functional analysis in different areas (biology and philosophy of mind) see, e.g., WRIGHT, 1973, CUMMINS, 1975, FODOR, 1985.

one of the apparently necessary components of factual decision-making. Remember what the question was: how can systems work without one of such elements, if we assume that assessment criteria and standards of proof are necessary components of legal fact-finding? The functional equivalence idea would explain the puzzle away because it would claim that criteria and standards perform the same function. So there would be no mystery in the fact that systems which have one but not the other component work fairly well. It would be unnecessary to have both, it would be redundant. Because they perform the same function, namely the function of guiding factual decision-making.

The functional equivalence idea is interesting but has some drawbacks. One is the (perhaps annoying) generality of the function appealed to: «guiding factual decision-making» is remarkable for its lack of specificity; but it's true that functional accounts always abstract from many details, insofar as they consider that such details are not relevant to the performance of the function envisaged (if the function is writing you don't bother about the differences between pens and pencils, which you bother about if you need devices which draw signs that cannot be canceled)⁴⁹.

A further and more significant drawback is the failure of the functional equivalence idea to make sense of some interesting cases. What about the Italian system of civil procedure? As stated, it has an explicit assessment criterion but no explicit standard. And there is an increasing debate about the standard or the standards that judges should use. If criteria and standards were functionally the same, this debate would be meaningless because the problem would not subsist. On the contrary there is a debate because it is unclear whether there is a standard and what this standard might be, whereas no one has doubts about the existence of an assessment criterion (Art. 116 of the civil procedure code) however vague it may be. The existence of this assessment criterion doesn't solve by itself the standard of proof problem. So there is no equivalence. Moreover, what about the Chilean system of criminal procedure? As mentioned, in this system you find both a standard (BARD) and an assessment criterion («sound criticism»). In this respect, assuming the functional equivalence approach, the system would be redundant. This conclusion would follow from the idea that the two perform the same function. It would be redundant to have a standard once you have a criterion, or vice versa. Now, this understanding of the Chilean system seems uncharitable, if not unreasonable, and I consider this to militate against the functional equivalence approach.

Indeed, Chile is not the only system that has both things explicit: the criminal procedure code of Colombia mandates BARD, since according to Art. 7 and Art. 381 guilt must be proven «más allá de toda duda» and according to Art. 372 evidence will serve to generate judicial knowledge of the relevant facts «más allá de duda razonable»⁵⁰. At

⁴⁹ On functional accounts of law see GREEN, 1998 and EHRENBURG, 2016. See also the sociological literature, e.g., PARSONS, 1951 and FRIEDMAN, 1975.

⁵⁰ See also Art. 91 and Art. 101 of the same code.

the same time the code mandates (i) a general holistic assessment criterion (according to Art. 380 evidentiary items will be assessed «en conjunto»)⁵¹ and (ii) more specific assessment criteria for specific kinds of evidence (for instance, Art. 404 states that lay testimony will be evaluated considering scientific and technical knowledge about perception and memory, checking in particular the nature of the thing perceived, the functioning of the witness' perceptual system, the circumstances of the perception, the memory processes and the demeanor of the witness at trial)⁵². All of this would be superfluous and redundant were standards and criteria functionally equivalent; on the contrary, such a system is highly interesting for the level of detail of the regulation concerned. The French criminal procedure system is not substantially different if *intime conviction* is taken as a standard (Art. 353 and Art. 427 of the criminal procedure code) and free assessment is the relevant criterion of assessment (Art. 428). And again, this would be redundant if criteria and standards were functionally equivalent.

A more promising idea is that of a functional connection between criteria and standards: *criteria indicate how to assess evidence and standards indicate how to make a decision once the evidence at disposal has been assessed*. A system like the Chilean code of criminal procedure makes everything explicit in requiring an assessment according to the principle of «sound criticism» and a decision according to the BARD standard. Similarly, for the even more detailed system of Colombia. The systems which are less explicit have just one component stated in positive law; then there are two possibilities: 1) that the other component is somehow incorporated in the explicit one; 2) that the other component is undetermined.

According to the first scenario, there can be explicit standards of proof and implicit assessment criteria, or vice versa explicit criteria and implicit standards. The incorporation idea consists in the fact that some system implicitly conveys a certain component (criterion or standard) incorporated in the explicit one or in the rest of the normative context (its rules, principles, etc.). I would locate the Italian criminal procedure system in this category, since it is characterized by an explicit standard (the BARD one) and an implicit assessment criterion consisting in a free but rational assessment of the evidence. Similarly, common law systems that adopt the BARD standard appear to convey at the same time the requirement of a reasonable assessment of the evidence; judicial instructions to juries perfectly fit this picture. Similarly, «legal proof» criteria usually incorporated standards of proof in that they set evidence thresholds and decision rules. Consider again the traditional two-witness rule, according to which two independent testimonies of the same amounted to «full

⁵¹ Similar holistic criteria can be found in European systems about circumstantial evidence: see, e.g., Art. 1382 of the French civil code and Art. 2729 of the Italian civil code on presumptions; cf. Art. 192 c. II of the Italian criminal procedure code.

⁵² In the same Colombian code see also, in particular, Art. 420 on scientific testimony and Art. 432 on documentary evidence. See also Art. 176 of the Colombian *Código general del proceso*: evidence shall be assessed «en conjunto» and complying with the «reglas de la sana crítica».

proof» of the fact: it was a rule about the value of testimonial evidence and, at the same time, it set a threshold for decision.

According to the second scenario, there can be systems where one of the necessary components of legal fact-finding is made explicit and the other is undetermined. What happens then in such systems? The analytical gap is filled by legal practice, with the possible support of legal theory and scholarly opinions. Legal practitioners and participants in legal disputes take sides about the undetermined component; arguments are developed and decisions are made; and in the long run, if no legislation intervenes, some component entrenched in legal practice will emerge as dominant through widely accepted arguments, judicial precedents, etc. The gap will be filled by legal practice. And in this kind of context the importance of legal argumentation is greater than elsewhere in this respect because legal practitioners and scholars need to construct and develop persuasive arguments to fill the gap. As an example, I would locate the Italian civil procedure system in this category, since it is characterized by an explicit assessment criterion and the lack of a positive standard of proof. This system differs from the criminal procedure system of Italy because in the latter, according to most commentators and practitioners, there is an implicit component (the implicit assessment criterion); as to the former, on the contrary, no general agreement exists about the component which is not explicit (the standard of proof); therefore, it is arguable that this component is undetermined and it is up to legal practice to fill in the gap. Recent judicial opinions in Italy confirm this reading.

But one may astutely object, saying the following: how could it be the case that Italian civil judges have been deciding disputes over the years if the standard of proof, being a necessary component of fact-finding, were completely undetermined? Decisions would have been impossible. So, there must have been a standard. This is true. However, this sort of *reductio ad absurdum* of the argument I provided is fallacious because it does not take into account the gap-filling role of practitioners, attorneys and judges in particular: these social actors contribute to that gap-filling even if what they do is not transparent to them. Even before they used the Anglo-American vocabulary and conceptual apparatus of standards of proof, for sure Italian legal practitioners had in their mind some idea about the evidence threshold that justified a civil decision, for it is plausible to think that they had some intuition about such sufficiency threshold. (Of course, the point applies to other jurisdictions.) When is the evidence sufficient to justify a certain belief about the litigated facts? When is it sufficient for the acceptance of a certain version of the litigated facts? You need not have the vocabulary of the «standards of proof» to grasp this idea and the problem. Then solutions come along with legal practice and argumentation.

So, in some way or other, both assessment criteria and standards of proof are needed. However, one may think that current legal theory is «forcing» the practice of civil law systems. It is putting pressure on legal practitioners, so the arguments would go, to compel them to use the standards vocabulary. To address this remark, recall some of the facts. Why don't civilians usually discuss standards? Because they do criteria. Why do common lawyers neglect assessment criteria? Because they focus

on standards. To have an explicit standard is, most of the time, to have an implicit criterion. And to have an explicit criterion is, most of the time, to have an implicit standard. Therefore, theory does not force practice if it makes explicit what is implicit in the system because of some functional connection. And theory does not force practice if it tries to give an accurate account of its gap-filling role when a necessary component of fact-finding is undetermined.

In sum, the functional connection account successfully explains how systems work without one of such analytically necessary components of factual decision-making (or, better, without one such component being made explicit in positive law). Systems work nevertheless either because there is some incorporation of the lacking element in the explicit one, or because legal practice fills the gap when one such component is undetermined.

5. CONCLUSION

After distinguishing assessment criteria from standards of proof, the first question we asked (in § 2) was why some jurists and courts do not see their difference. The second question (in § 4) was why some systems have criteria and others have standards. And the third question (again in § 4) was how systems work without one of such necessary components of legal fact-finding.

The answer to the first question came with the answer to the second. Some historical, institutional and procedural differences explain why some systems have assessment criteria and others have standards. This is basically the answer to the second question. And this answer suggests that jurists belonging to one context are prone to see also elsewhere what they are more familiar with. This means that common lawyers have some tendency to take assessment criteria as standards and, vice versa, civil lawyers have some tendency to see standards as assessment criteria. The less familiar thing is understood in terms of the more familiar one. More importantly, when different systems come to overlap to a certain extent (as is the case in the jurisdiction of courts like the ECtHR) parties and judges who come from different contexts and backgrounds run the risk of mixing things up. Some terms, concepts and institutions that belong to one context are used in combination with terms, concepts and institutions of a different context. Some merging of the issues (assessment and standards) comes from the work of supranational courts such as the ECtHR, insofar as they use some traditional common law terminology and conceptual apparatus to deal with civil law countries issues, or vice versa. Not surprisingly, the result can be confused and confusing to a certain extent. This is one of the prices of a more integrated or even globalized legal world, where different concepts and institutions play together and interact in ways that are not always foreseeable⁵³.

⁵³ For a similar point see HANS & GERMAIN, 2011: 738, 761-763. Cf. DAMAŠKA, 1997b.

On the other hand, there are international jurisdictions that set things more clearly: for instance, the Rome Statute of the International Criminal Court (ICC) mandates BARD as the conviction standard of proof (Art. 66.3), while the related Rules of Procedure and Evidence establish the free assessment of evidence (Rule 63.2) for relevance and admissibility purposes (in accordance with the discretion conferred upon the decision-makers by Art. 64.9, Art. 69.4 and Art. 74.2 of the Rome Statute)⁵⁴. Here standards and criteria are clearly distinguished. Also, the International Court of Justice has a clear view; for instance, in the well-known 2007 decision on *Bosnia and Herzegovina v. Serbia and Montenegro*, the Court is careful to distinguish three «matters»: the burden of proof (§§ 204-207), the standard of proof (§§ 208-210) and the methods of proof (§ 211ff) as assessment of evidentiary weight (§ 213).

The answer to the third question asked above is more theoretically intriguing. It consists in the claim that there is a functional connection between assessment criteria and standards of proof. Because of that connection, one element can incorporate the other. And because of that connection, where one of those things is neither made explicit in positive law nor incorporated in the explicit component, legal practice and scholarly literature make an effort to fill in the relevant gap, providing answers and solutions to the questions and problems that actual cases pose to legal practitioners and scholars.

6. BIBLIOGRAPHY

- ACCATINO, D., 2011: «Certezas, dudas y propuestas en torno al estándar de la prueba penal», *Revista de Derecho de la Pontificia Universidad Católica de Valparaíso*, 37(2): 483-511.
- ALLEN, R.J., 1988: «Idealization and Caricature in Comparative Scholarship», *Northwestern University Law Review*, 82(3): 785-807.
- 1991: «The Nature of Juridical Proof», *Cardozo Law Review*, 13: 373-422.
- 2014: «Burdens of Proof», *Law, Probability and Risk*, 13: 195-219.
- ALLEN, R.J. et al., 1988: «The German Advantage in Civil Procedure: A Plea for More Details and Fewer Generalities in Comparative Scholarship», *Northwestern University Law Review*, 82(3): 705-762.
- ALLEN, R.J. and PARDO, M.S., 2007: «The Problematic Value of Mathematical Models of Evidence», *The Journal of Legal Studies*, 36(1): 107-140.
- 2019: «Relative Plausibility and Its Critics», *The International Journal of Evidence and Proof*, 23(1-2): 5-59.
- ALLEN, R.J. and STEIN, A., 2013: «Evidence, Probability, and the Burden of Proof», *Arizona Law Review*, 55: 557-602.
- ANDERSON, T., SCHUM, D. & TWINING, W., 2005: *Analysis of Evidence*, 2nd ed., Cambridge: Cambridge University Press.
- BARZUN, C., 2008: «Rules of Weight», *Notre Dame Law Review*, 83(5): 1957-2017.

⁵⁴ See CUMIZ, 2016 for a discussion on BARD and a reference to the other standards of proof used by the ICC for different procedural purposes. But see also, for a critical account of international criminal trials, COMBS, 2010. Cf. JACKSON & SUMMERS, 2012: 108 ff and GASKINS, 2016.

- Bayón, J.C., 2008: «Epistemología, moral y prueba de los hechos: hacia un enfoque no benthamiano», *Analisi e diritto*, 2008: 15-34.
- BERGEAUD, A., 2010: *Le droit à la preuve*, Paris: L.G.D.J.
- BELL, R.S., 1987: «Decision Theory and Due Process: A Critique of the Supreme Court's Lawmaking for Burdens of Proof», *Journal of Criminal Law and Criminology*, 78: 557-585.
- BOHLANDER, M., 2012: *Principles of German Criminal Procedure*, Oxford: Hart Publishing.
- BREDIN, J.-D., 1996: «Le doute et l'intime conviction», *Droits*, 23: 21-29.
- BOULOC, B., 2004: «La preuve en matière pénale», in PUIGELIER, C. (ed.), *La preuve*, Paris: Economica: 43-56.
- BOVINO, A., 2005: «Evidential Issues before the Inter-American Court of Human Rights», *Sur. International Journal on Human Rights*, 2: 56-79.
- CAPRIOLI, F., 2009: «L'accertamento della responsabilità penale "oltre ogni ragionevole dubbio"», *Rivista italiana di diritto e procedura penale*, 52: 51-92.
- CARTIER, M.-E., 2004: «Brèves remarques sur la preuve devant la Cour pénale internationale», in PUIGELIER, C. (ed.), *La preuve*, Paris: Economica: 57-72.
- CASTILLO DE LA TORRE, F., 2009: «Evidence, Proof and Judicial Review in Cartel Cases», *World Competition*, 32(4): 505-578.
- CLERMONT, K. & SHERWIN, E., 2002: «A Comparative View of Standards of Proof», *The American Journal of Comparative Law*, 50: 243-275.
- COHEN, L.J., 1977: *The Probable and the Provable*, Oxford: Clarendon Press.
- COLOMA, R., 2012: «¿Realmente importa la sana crítica?», *Revista Chilena de Derecho*, 39(3): 753-781.
- COLOMA, R. and AGÜERO, C., 2014: «Lógica, ciencia y experiencia en la valoración de la prueba», *Revista Chilena de Derecho*, 41: 673-703.
- COMBS, N.A., 2010: *Fact-Finding Without Facts. The Uncertain Evidentiary Foundations of International Criminal Convictions*, Cambridge: Cambridge University Press.
- CUMIZ, J.A., 2016: «El estándar de prueba "más allá de toda duda razonable" en la jurisprudencia de la Corte Penal Internacional», *Revista de Derecho Penal y Procesal Penal*, 2016/1: 5-21.
- CUMMINS, R., 1975: «Functional Analysis», *The Journal of Philosophy*, 72: 741-765.
- DAHLMAN, C., 2015: «The Felony Fallacy», *Law, Probability and Risk*, 14: 229-241.
- DAMAŠKA, M., 1986: *The Faces of Justice and State Authority. A Comparative Approach to the Legal Process*, New Haven and Chicago: Yale University Press.
- 1997a: *Evidence Law Adrift*, New Haven & London: Yale University Press.
- 1997b: «The Uncertain Fate of Evidentiary Transplants: Anglo-American and Continental Experiments», *The American Journal of Comparative Law*, 45: 839-852.
- 2003: «Epistemology and Legal Regulation of Proof», *Law, Probability and Risk*, 2: 117-130.
- 2018: *Evaluation of Evidence. Pre-Modern and Modern Approaches*, Cambridge: Cambridge University Press.
- DELLA TORRE, J., 2014: «Il lungo cammino della giurisprudenza italiana sull'«oltre ogni ragionevole dubbio"», *Diritto Penale Contemporaneo* (<https://www.penalecontemporaneo.it>), June 20, 2014.
- DELMAS-MARTY, M., 1996: «La preuve pénale», *Droits*, 23: 53-65.
- EHRENBERG, K.M., 2016: *The Functions of Law*, Oxford: Oxford University Press.
- FERRAJOLI, L., 1989: *Diritto e ragione. Teoria del garantismo penale*, Roma-Bari: Laterza.
- FERRER, J., 2007: *La valoración racional de la prueba*, Madrid: Marcial Pons.
- FODOR, J., 1985: «Fodor's Guide to Mental Representation», *Mind*, 94: 76-100.
- FRIEDMAN, L.M., 1975: *The Legal System. A Social Science Perspective*, New York: Russell Sage.
- GASCÓN, M., 2010: *Los hechos en el derecho. Bases argumentales de la prueba*, 3rd ed., Madrid: Marcial Pons.
- GASKINS, R., 2016: «The Legal Characterization of Facts at the International Criminal Court», in FETERIS, E. et al. (eds.), *Legal Argumentation and the Rule of Law*, The Hague: Eleven International Publishing: 59-68.
- GOLDMAN, A., 2002: «Quasi-Objective Bayesianism and Legal Evidence», *Jurimetrics*, 42: 237-260.

- GONZÁLEZ, J., 2006: «La fundamentación de las sentencias y la sana crítica», *Revista Chilena de Derecho*, 33(1): 93-107.
- GONZÁLEZ LAGIER, D., 2013: *Quaestio facti: Ensayos sobre prueba, causalidad y acción*, Mexico: Fontamara.
- GREEN, L., 1998: «The Functions of Law», *Cogito*, 12(2): 117-124.
- HAACK, S., 2014: *Evidence Matters. Science, Proof, and Truth in the Law*, Cambridge: Cambridge University Press.
- HANS, V. & GERMAIN, C., 2011: «The French Jury at a Crossroads», *Chicago-Kent Law Review*, 86(2): 737-768.
- IACOVIELLO, F.M., 2006: «Lo standard probatorio dell'al di là di ogni ragionevole dubbio e il suo controllo in Cassazione», *Cassazione penale*, 46: 3869-3884.
- JACKSON, J.D. & SUMMERS, S.J., 2012: *The Internationalisation of Criminal Evidence. Beyond the Common Law and Civil Law Traditions*, Cambridge: Cambridge University Press.
- KAPLOW, L., 2012: «Burden of Proof», *The Yale Law Journal*, 121: 738-859.
- LANGBEIN, J.H., 1985: «The German Advantage in Civil Procedure», *The University of Chicago Law Review*, 52(4): 823-866.
- LANGBEIN, J.H., 1988: «Trashing “The German Advantage”», *Northwestern University Law Review*, 82(3): 763-784.
- LAUDAN, L., 2005: «Por qué un estándar de prueba subjetivo y ambiguo no es un estándar», *Doxa*, 28: 95-113.
- 2006: *Truth, Error, and Criminal Law. An Essay in Legal Epistemology*, Cambridge: Cambridge University Press.
- LILLQUIST, E., 2002: «Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability», *U. C. Davis Law Review*, 36: 85-197.
- LLUCH, X.A., 2012: «La dosis de la prueba: entre el *common law* y el *civil law*», *Doxa*, 35: 173-200.
- MACKOR, A.R., 2016: «Improving Judicial Argumentation about Evidence in Dutch Unus Testis Cases», in FETERIS, E. et al. (eds.), *Legal Argumentation and the Rule of Law*, The Hague: Eleven International Publishing: 111-122.
- NOVAK, M., 2016: «Arguing for Certainty in Criminal Decision-Making and the Rule of Law», in FETERIS, E. et al. (eds.), *Legal Argumentation and the Rule of Law*, The Hague: Eleven International Publishing: 147-157.
- PADOA-SCHIOPPA, A. (ed.), 1987: *The Trial Jury in England, France, Germany 1700-1900*, Berlin: Duncker & Humblot.
- 1994, *La giuria penale in Francia*, Milano: LED.
- 2003, *Italia ed Europa nella storia del diritto*, Bologna: il Mulino.
- PARSONS, T., 1951: *The Social System*, Glencoe: The Free Press.
- PAÚL, Á., 2012: «In Search of the Standards of Proof Applied by the Inter-American Court of Human Rights», *Revista Instituto Interamericano de Derechos Humanos*, 55: 57-102.
- PICINALI, F., 2013: «Two Meanings of “Reasonableness”: Dispelling the “Floating” Reasonable Doubt», *The Modern Law Review*, 76(5): 845-875.
- PUTNAM, H., 1981: *Reason, Truth and History*, Cambridge: Cambridge University Press.
- 1995: «Are Moral and Legal Values Made or Discovered?», *Legal Theory*, 1: 5-19.
- REDMAYNE, M., 1999: «Standards of Proof in Civil Litigation», *The Modern Law Review*, 62: 167-195.
- 2003: «Objective Probability and the Assessment of Evidence», *Law, Probability and Risk*, 2: 275-294.
- ROBERTS, P. & ZUCKERMAN, A., 2010: *Criminal Evidence*, 2nd ed., Oxford: Oxford University Press.
- SCHWEIZER, M., 2014: «Comparing Holistic and Atomistic Evaluation of Evidence», *Law, Probability and Risk*, 13: 65-89.
- 2016: «The Civil Standard of Proof—What Is It, Actually?», *The International Journal of Evidence and Proof*, 20(3): 217-234.
- STEIN, A., 2005: *Foundations of Evidence Law*, Oxford: Oxford University Press.
- TARUFFO, M., 1992: *La prova dei fatti giuridici*, Milano: Giuffrè.

- 2003a: «Rethinking the Standards of Proof», *The American Journal of Comparative Law*, 51: 659-677.
- 2003b: «Algunos comentarios sobre la valoración de la prueba», *Discusiones*, 3: 81-97.
- 2009: *La semplice verità. Il giudice e la costruzione dei fatti*, Roma-Bari: Laterza.
- TILLERS, P. & GREEN, E.D. (eds.), 1988: *Probability and Inference in the Law of Evidence*, Dordrecht: Kluwer Academic Publishers.
- TRIBE, L.H., 1971: «Trial by Mathematics. Precision and Ritual in the Legal Process», *Harvard Law Review*, 84(6): 1329-1393.
- TUZET, G., 2020: «Assessment Criteria or Standards of Proof? An Effort in Clarification», *Artificial Intelligence and Law*, 28(1): 91-109.
- UBERTIS, G., 2013: «Razionalismo processuale e verità», *Il giusto processo civile*, 1/13: 325-333.
- VÁZQUEZ, C. (ed.), 2013: *Estándares de prueba y prueba científica: ensayos de epistemología jurídica*, Madrid: Marcial Pons.
- VERHEIJ, B. et al., 2016: «Arguments, Scenarios and Probabilities: Connections Between Three Normative Frameworks for Evidential Reasoning», *Law, Probability and Risk*, 15: 35-70.
- WHITMAN, J., 2008: *The Origins of Reasonable Doubt. Theological Roots of the Criminal Trial*, New Haven: Yale University Press.
- WRIGHT, L., 1973: «Functions», *The Philosophical Review*, 82: 139-168.

