Published in: Ferrer Beltrán, Jordi; Vázquez, Carmen (Eds.). Evidential Legal Reasoning: Crossing Civil Law and Common Law Traditions. Cambridge: Cambridge University Press. ISBN 978-1-316-51699-7, pp. 375-394. https://doi.org/10.1017/9781009032049.018

Is it possible to formulate a precise, objective standard of proof? Some questions based on an argumentative approach to evidence.

Daniel González Lagier
University of Alicante.

1. Introduction.

In the Latin American context, evidence theory has progressed in a few years from being a subject hardly dealt with by jurists and philosophers to become a highly active, flourishing, thriving field of study, with important advances and intriguing proposals (this development occurred long beforehand in the English-speaking world). However, I believe that, as in any expanding discipline, in order to have a solid base it is necessary to have precise terminology that does not result in serious errors, and to share a series of concepts allowing us to clearly formulate problems and avoid confusion due, not to the difficulty of the problems dealt with but to the lack of univocality in the language used. It is in relation to this aspect that I think it is possible and necessary to make further progress, particularly in relation to what might be considered one of the central planks of this discipline: the evaluation of evidence and the decision on which hypothesis must be accepted as proven. Our legal systems normally use quite similar terminology to refer to the laws of logic, the lessons of experience, scientific knowledge and the rules of reasoned judgment as evaluation criteria, but without specifying these rather vague concepts. To these criteria would be added the standard of proof, the subject of so many discussions in our recent literature. However, in my opinion, despite these discussions, the relationships and differences between the standard of proof and the evaluation criteria remain unclear. This lack of precision can make it difficult to discuss which evaluation criteria and which particular standards are appropriate, or even to discuss the possibility of formulating a precise, objective standard of proof. In this study, I will try to offer a set of conceptual suggestions that could be used to make progress in the search for a shared terminological and conceptual basis on this point. To do this, I will adopt an argumentative perspective on evidence, focusing on three points: (1) The structure of evidentiary inference; (2) which reasons count as good ones for establishing the degree of corroboration of a hypothesis and (3) the possibility of formulating a precise, objective standard of proof.

2. Three ways of arguing about facts.

One of the many senses in which the word "evidence" is used is "evidence as argument".[1] From this point of view, "proving" something consists of constructing an argument to justify a certain hypothesis as the factual premise for a legal decision. As I see it, this kind of argument always consists of correlating two types of facts (or statements about facts): the facts we want to prove and the facts we use to prove them (the elements of judgment). This argument therefore consists of a set of premises (the elements of judgment); a conclusion (the hypothesis on the facts we want to prove); and a connection or relationship between the premises and the hypothesis. This link or connection between the elements of judgment and the hypothesis may be empirical, normative or conceptual.

In the first case, the link is an empirical generalisation correlating facts like those described in the premises with facts like those described in the conclusion, based on the observation of a past association between the two types of facts. These generalisations can include scientific knowledge, and we can call them "maxims of experience", although sometimes this term refers only to the empirical generalisations attempted based on common sense and general acceptance. In these cases, we can speak of empirical evidential inference (in some contexts it would be appropriate to restrict the word "evidence" to these circumstances).

For example:

1. The defendant was arrested near the house where the burglary was committed shortly after the time that it happened, carrying the objects removed from the house together with a metal lever (elements of judgment).
2. If someone is surprised at the scene of the crime or nearby immediately after the event carrying items or the proceeds of the crime and/or items necessary for committing it, they are probably responsible for the crime (lesson of experience).
3. The defendant is responsible for the crime (proven fact).

In the second case, the link is a rule (normally from legislation or jurisprudence) establishing that, if there are facts like those described in the premises (the basic fact), a certain hypothesis must be considered proven (the consequential fact). We can call these rules the "rules of presumption" or the rules of the weighted evaluation of evidence and these inferences are known as normative evidential inferences.

For example:

1. Subjects *x* and *y* (father and son) died in the same car accident and there is no evidence of who died first (elements of judgment).

[1] Daniel González Lagier: "Presunción de inocencia, verdad y objetividad", in GARCÍA AMADO (coord.), *Prueba y razonamiento probatorio en Derecho*, Granada, 2014, pp. 109 and ff.

2. “If there is any doubt as to which of two or more people who succeed one another died first (...), and if there is no proof, they are presumed to have died at the same time” (legal presumption established in article 33 of the Spanish Civil Code).
3. *x* and *y* died at the same time (proven fact).

In the third case, the connection is established by a conceptual definition or rule establishing that the facts of the kind described in the hypothesis “count as” (in other words, “can be subsumed in”) a certain category of facts (an action, an intention, a causal relationship, etc.). In these cases, what is at stake is not so much whether a particular fact has occurred but rather its interpretation; in other words, its classification in a particular general category of facts.

For example:

1. Everyone who contracted the toxic syndrome had eaten rape-seed oil, but not everyone who ate rape-seed oil contracted the toxic syndrome. In other words, rape seed oil was a necessary but not sufficient condition of the toxic syndrome (elements of judgment).
2. When one fact is a necessary condition of another (even if it is not necessary and sufficient) the former is a cause of the latter (definition).
3. The rape-seed oil caused the toxic syndrome (proven fact).

In all cases, as might be expected, empirical evidential inferences are prioritary. It is impossible to make one of the other types of inference without proving that the basic fact of the presumption or the definition has occurred (which will have to be done by using an empirical generalisation). In the following discussion, I will leave aside evidential inferences based on definitions, which raise different difficulties to the ones I want to deal with here[2].

3. Concerning the evaluation of evidence and the confirmation of a hypothesis

3.1. Evaluating evidence:

It is clear that the type of argument we call “evaluation of evidence” occurs only as a result of empirical evidential inference. In the case of normative evidential inference – in other words in cases of weighted evidence – the evaluation is already predetermined in the rule[3]. Inferences based on empirical generalisations correspond to free evidence evaluation systems, while normative evidential inferences belong to weighted evidence systems. A “perfect” free evidence evaluation system would have no rules for establishing presumptions. By contrast, a “perfect” weighted evidence system is clearly impossible unless it is

[2] On these, see Daniel González Lagier, “Hechos y conceptos”, *Cuadernos Electrónicos de Filosofía del Derecho*, no. 15.
[3] When we use a rule of presumption, we will have to prove the basic fact of the presumption and also (if accepted) any possible evidence in rebuttal, but this will be done by empirical evidential inference.

completely circular, because at some point the basic fact of a rule of presumption must be empirically proven. In fact, although our systems are considered to be free evidence evaluation systems, within them there are cases of "free evidence" (or freer evidence) and cases of "weighted evidence" (or less free evidence). The use of free evaluation or weighted evidence is a question of degree.

It is, then, in cases where the judge is free to examine whether the elements of judgment make it possible to support the hypothesis and to what degree this can be done that we can properly speak of the evaluation of evidence. However, empirical evidential inference does not allow an absolutely certain conclusion to be drawn. On the contrary, it allows us to know the truth only in a limited and rather approximate way. This is the case even if we formulate the inference as a deduction, because it is not possible for us to be more certain of the conclusion than we are of the premises – we must not confuse the logical validity of the argument with the material certainty of its conclusion. From the point of view of argument, the evaluation of evidence can be identified with the degree to which the empirical evidential inference is correct or solid; in other words, the degree to which the evidence confirms or corroborates the hypothesis. We might also say that the evaluation of evidence consists of determining the level of inductive probability with which the hypothesis/conclusion follows from the premises (in other words, from the elements of judgment and the lesson of experience). We therefore need rational criteria to determine the degree to which the conclusion is solid. These criteria are not formal, or are not only formal. Formal criteria would be the logical rules also alluded to by our systems as evaluation criteria. Seen from the point of view of argument, the rules of reasoned judgment can be interpreted as informal criteria for the solidity of empirical evidential inference.

3.2. The "rules of reasoned judgment".

In previous studies I have suggested the following criteria or rules for rational evidence evaluation[4]:

1) The more elements of judgment we have in favour of a hypothesis, the better the confirmation of the hypothesis.

2) The more varied the elements of judgment (in other words, if they add information making it possible to eliminate alternative hypotheses), the better the confirmation of the hypothesis.

3) The more relevant the elements of judgment (the better related they are to the hypotheses through reliable empirical generalisations), the better the confirmation of the hypothesis.

[4] Daniel González Lagier, *Quaestio facti. Ensayos sobre prueba, causalidad y acción*, Editorial Palestra-Temis, Lima, 2005.

4) The more reliable the elements of judgment (the better founded they are in other elements of judgment and previous inferences or direct observations or firm knowledge), the better the confirmation of the hypothesis.
5) The better founded the lessons of experience in inductive generalisations, the firmer the hypothesis.
6) The greater the probability expressed in the lesson of experience, the firmer the hypothesis. (Lessons of experience have the following structure: “If p, then probably q”; the degree of probability with which the two types of fact are correlated is important for the confirmation of the hypothesis).
7) The hypothesis must not have been refuted either directly (no fact incompatible with the hypothesis must have been proved) or indirectly (hypotheses that are true must not be refuted if the truth of the main hypothesis is accepted).
8) If the hypotheses derived from the main hypothesis (in other words, hypotheses that would be true if the main hypothesis was true) can be confirmed, the better the confirmation of the main hypothesis (by means of abductive argument).
9) The more coherent the hypothesis from a narrative point of view, the better the confirmation of the hypothesis.
10) The more elements of judgment explained by the hypothesis, the better the confirmation of the hypothesis.
11) The fewer unproven facts required for the hypothesis to be true, the better the confirmation of the hypothesis.
12) The fewer existing alternative hypotheses incompatible with the main hypothesis, the better the confirmation of the main hypothesis.

I believe it is enlightening to identify the “rules of reasoned judgment” to which our systems allude with criteria of epistemological rationality like these. There is room for discussion on many of my proposed points: there may be a rule missing; they may not be well formulated; some of the rules are redundant (2 and 12, for example, are the same rule seen from the point of view of elements of judgment and from the point of view of the hypothesis, while 3 and 5 also point towards the same idea, from the point of view of elements of judgment and the point of view of the lesson of experience); some of them may be superfluous, incorrect, or defective in other ways. They can probably be presented more clearly, economically and precisely. However, what I want to suggest is that what in our culture we call rules of reasoned judgment must be rules of this kind if the aim is for the evaluation of evidence to be epistemologically rational (and, therefore, tend to ensure conclusions that are probably true, or that minimise error). They certainly cannot differ very much from them.

Does it make sense for legislators or jurisprudence to regulate or positivise this kind of rule? I believe it is important to realise that, whether or not they are positivised, these methodological rules are necessary in terms of trying to infer rationally correct hypotheses based on the available elements of judgment. Moreover, the fact that this is necessary does not depend directly on judicial authority. It does indirectly, however, in that the design and purpose of the process of proof depends on the judicial authorities. In other words, what the Law establishes as *compulsory* is the requirement of rationality in the evaluation of evidence; the evaluation rules are *necessary* means to this end. Just as the logical principle of non-contradiction must be respected, whether the legislator spells this out or not, the intrinsic nature of the laws of logic does not change. Nor do the rules of epistemological rationality need to be positivised to make them binding or necessary. They would only be included in normative texts, therefore, in the form of guidelines, examples and indications. The question of what the rules of epistemological rationality are is, in itself, a methodological and philosophical issue open to discussion and dependent on the epistemological theory assumed. It is therefore not advisable to positivise it, except perhaps for the more flexible jurisprudential route.

One clarification: I believe it is important to point out that, while the evaluation of evidence is identified with estimating the degree of solidity of evidential inference, the subject of evaluation is not merely the *evidence* (the elements of judgment) but rather the *proof* – the evidential argument as a whole. In other words, as we have seen, it concerns the criteria covering the elements of judgment, the hypothesis and the connection between these.

3.3. Rules of reasoned judgment and lessons of experience:

The rules of reasoned judgment are different from the lessons of experience. The former are normative (although not judicial – as we have just said they are requirements for rationality determining which forms of argument are correct and which are not, and they can be seen as a set of rules determining the framework of theoretical rationality). Their basis is also not empirical (unless some type of naturalised epistemology is maintained, like that proposed by Quine, for whom epistemology must be reduced to cognitive psychology).[5] A rule establishing that "the more elements of judgment there are in favour of a hypothesis, the firmer it is" is not something we could ultimately justify with experience (although it is possible that we might have learned from observing the criteria used by others). Trying to justify this kind of rule of inductive rationality from experience would raise various problems. If we try to show that these rules are successful in finding the truth there would be a problem

[5] W.V. Quine, "Naturalización de la epistemología", in *La relatividad ontológica y otros ensayos*, Ed. Tecnos, 2002.

of circularity, because to show that they are normally successful we would have to use the principles or rules that we are trying to justify; if we simply try to justify them by showing that they are in fact the criteria used by the majority to justify beliefs, they would also lose their normative dimension (as they would be no use in determining whether or not an argument is correct, they would simply indicate whether or not an argument conforms to a habit). Meanwhile, the lessons of experience are empirically-based descriptive statements (and therefore either true or false). We arrive at them through general argument (using the rules of epistemological rationality) based on examining particular cases. These are necessary to correlate evidentiary facts and the facts that require proof, but it is not a logical or inductive need: they are required as premises of evidentiary inference, not as methodological rationality criteria.

3.4. The gradual nature of the confirmation of hypotheses.

The rules for evaluating evidence are gradual in at least two senses: firstly concerning the criteria, which are themselves gradual (varying number of elements of judgment, varying degree of reliability, varying degree of coherence of hypotheses, varying degree of foundation of the maxims of experience and so on). Secondly, one hypothesis could be justified by several rules, and it would be too demanding to require that it should meet all of them to a relevant degree.

There are two consequences of this gradual nature. The first, which establishes the degree of confirmation of a hypothesis, requires an overall judgment in the light of all these criteria, which means it is necessarily the result of a holistic evaluation.

None of these criteria alone is a necessary or a sufficient condition of a certain degree of confirmation. They are not a necessary condition because if one of them is absent it can always be made up for by other criteria. For example, it cannot be said that, as the number of elements of judgment in favour of a hypothesis is very small, the hypothesis necessarily has a very low degree of confirmation, because this could be compensated for by the fact that the lessons of experience connecting these elements of judgment with the hypothesis are very solid and that it has been possible to eliminate a good number of alternative hypotheses. Nor is it a sufficient condition for a certain degree of confirmation, because the criteria it has in its favour can always be counteracted by a deficit in others. For example, it cannot be said that the fact that the lessons of experience are very well-founded guarantees that the hypothesis is properly confirmed, because this could be undermined by the fact that the elements of judgment themselves are not very reliable. As we will see, this is important in the discussion about the plausibility of a standard of proof used as a sufficiency threshold.

The second consequence is that these criteria make it possible to determine the relative probability (in logical or inductive terms) of one hypothesis compared with another, but not how much more probable it is.[6] *In other words, they allow the comparison and ordering of the degree of justification of different particular hypotheses but not a numerical quantification of their probability. So now, once the evidence has been evaluated, the problem of making the decision arises: is the degree of confirmation obtained sufficient to consider the hypothesis proven? Answering this question requires a new criterion: the standard of proof. This criterion must operate as a threshold (although a certain degree of vagueness is acceptable) allowing us to discriminate between what we consider to be (sufficiently) proven and what we do not consider to be (sufficiently) proven.*

Therefore, the evaluation criteria and the standard of proof have different objectives and purposes: what we evaluated using the evidence evaluation rules is the evidentiary argument or inference to try to establish their degree of confirmation or justification. Meanwhile, the standard of proof does not attempt to evaluate the evidence against. Instead, we use it to evaluate the degree of justification obtained (in other words, the result of previous argument) to answer the question of whether it is sufficient for the decision to be made. Of course, we do not need just one standard of proof, it may also be different – more or less strict – depending on the type of decision involved. However, the evaluation criteria are the same for all cases (although some may be more relevant or more often used for some types of facts).

4. Standards of proof.

4.1. Practical standards and decisions

Standards of proof (or decision) are not exclusively a judicial problem. As we know, our decisions and actions can be seen as the result of the combination of a desire and a belief about how to satisfy it. That means our knowledge of the world (our beliefs) has practical and not just theoretical relevance. We need beliefs to know how to act. However, the consequences of our decisions and actions may be relevant to different degrees. The more far-reaching the decision I have to make, the more serious its consequences, and the more certain I need to be of the beliefs guiding that decision. If my life depends on arriving in Madrid on time, my belief that the plane leaves at nine in the morning based only on what I have read in the paper seems not to be sufficiently justified. The reasonable thing to do would be to try to reach a higher level of certainty, so I should look for more evidence. The level to which we demand that a belief should be justified depends on the context and the

[6] For an influential analysis of the notion of inductive probability, see L.J. Cohen, *The probable and the provable*, Oxford: Clarendon Press, 1977.

practical relevance of the belief. So the same belief with the same inductive support may or may not be sufficiently justified depending on the context. For example:[7]

- Context 1: On Friday, Michael and his wife go to the bank to pay money in. As there is a long queue, Michael says: "I'll come back on my own tomorrow." His wife says: "Perhaps the bank isn't open tomorrow. A lot of banks are closed on Saturdays." Michael answers: "No. I know it will be open: I came on Saturday two weeks ago and they're open until lunchtime."
- Context 2: This time they need to pay the money in before Monday as some cheques they have signed will be drawn on that day. If the funds are not paid in by Monday, they are going to have real problems. As in context 1, there is a long queue and Michael says he will come back the next day. His wife reminds him that, if they do not pay the money in before Monday, they are going to have difficulties and tells him: "Banks change their opening times. Are you sure the bank is open tomorrow?" Michael, who is just as convinced as before that the bank is open on Saturday, replies: "Well, no. We'd better stay and pay the money in today."

In context 1, Michael says he knows the bank is open on Saturdays; in context 2 he says he doesn't. The evidence in favour of his belief is the same in both cases. This evidence is sufficient to consider his belief justified in 1 but not to consider it justified in 2. From this, the philosophical view known as contextualism draws the conclusion that the attribution of knowledge (justified true beliefs) is sensible in the context. However, another way of looking at this relativity of the degree to which beliefs are justified with respect to the context, which avoids certain problems of contextualism, can be to introduce the distinction between *belief* and *acceptance* as two different types of propositional attitudes. Using this distinction, it should be said that it is not the case that in context 1 the belief is justified and in 2 it is not. In both cases the belief is equally justified, but this degree of justification in context 1 is sufficient for it to be accepted (to act in accordance with it), while in context 2 it is not.

4.2. Belief and acceptance.

I cannot *believe* something I know to be false; by contrast I can *accept* something even though I have doubts, or even if I believe it to be false, and act as if it were true. Acceptance is therefore a propositional attitude which is also related to truth, but in a different way to belief. A person who believes something considers that their belief is true, but this consideration is not necessarily present in acceptance. A person who accepts something can only consider that there are reasons to act as if the statement were true, even if it is not. According to L.J. Cohen, "Accepting *p* means having or adopting a policy of judging, suggesting or postulating *p* – that is, including this proposition or rule among a one's own premises for deciding what to do or think in a particular context, whether or not *p* is

[7] Tobies Grimaltos, "Creencia, aceptación y conocimiento", *Episteme*, Vol. 29, No. 1, 2009, pp. 35-50.

actually true."[8] We can sketch out the differences between belief and acceptance in the following way:

a) Belief is gradual (we can be convinced of something to different degrees): acceptance is all or nothing.

b) Belief is determined by epistemic reasons – reasons for believing – but not by practical reasons (that it is wise to believe p is not a suitable reason to make one believe p). Acceptance is determined by epistemic and/or practical reasons.

c) Belief is not an action; in other words, it is not entirely within our control. (Beliefs can be consequences of our actions but they are not actions in themselves. For example, we cannot cease to believe *p*, for which we have overwhelming evidence, although if we do not yet believe *p* we can avoid looking for this evidence. And we cannot force ourselves to believe *p* if we do not have epistemic reasons for it, although we can look for evidence). Acceptance, on the other hand, is the result of a deliberate decision.

It follows from the above, then, that belief is a reason for acceptance, but it is not the only one. Acceptance can occur for epistemological or other reasons (prudence, for example). However, when we accept something for epistemological reasons, a certain degree of justification is necessary, and that degree of justification is a practical criterion related to the context and the purposes of the agent.

4.3. Standard of proof and acceptance.

The concept of acceptance can be useful to take account of some of the propositional attitudes present in the evidence. For example, when we argue through normative evidentiary inference (in other words, when we do not evaluate the evidence but simply subsume the elements of judgment in the factual situation of a rule establishing a presumption or determining an evidentiary result), it makes no sense pretending that rules can force us to have a particular belief, as beliefs are not entirely under our control. However, if acceptance is a deliberate action, rules can force us to accept a particular hypothesis as the factual premise for a judicial decision.[9] Some of these rules require acceptance of a particular evidential result for epistemic reasons (they force us to accept something because there are reasons to believe it). Others require acceptance for practical reasons (protection of a legal asset). In weighted evidence systems, then, the judge is required to accept the proven facts without wondering whether there are reasons to believe them.

The idea of acceptance can also shed some light on standards of proof. We must remember that the evaluation of evidence makes sense only in the case of empirical

[8] L.J. Cohen, *Belief and Acceptance*, Oxford: Clarendon Press, 1992, p. 4.
[9] Daniel Mendonca, "Presunciones", in *Doxa*, no. 28, 1998.

evidentiary inferences. Similarly, the problem of standard of proof arises only in this type of inference. A rule that establishes the obligation to consider a fact proven if a certain combination of elements of judgment are present (which we have called a presumption rule) establishes, as we have seen, that it must be accepted that this fact has occurred. That in itself is an acceptance criterion, so no new criterion – or standard – is required in order to know whether the fact has to be accepted. In legal or weighted evidence systems, the problem of determining the standard of proof does not arise (or, we might say, the legislator has already established a rigid standard of proof for each case). But in cases of free evidence evaluation we need criteria to tell us the degree to which the hypothesis must be justified so that it has to be accepted (or used as a guide for the judicial decision). In these free evaluation systems, there are two types of reasons for accepting the guilty hypothesis: firstly there are reasons for believing it in order to reduce error, which are those indicated by the evidence evaluation criteria. Secondly, there are what we might call secondary reasons for considering the degree of certainty or justification achieved – in other words reasons for acceptance. These secondary reasons are practical ones, related to how we want to distribute the cost of mistakes. In criminal law, for example, it is assumed that it is more serious to find an innocent person guilty than to acquit a guilty one, so the degree of sufficiency demanded must be higher. In other words, the hypothesis accepted must be epistemologically founded, but with a degree of justification that must exceed a certain threshold or meet certain requirements. For this reason, it is possible that, in cases of free evaluation, the judge believes the hypothesis is correct but does not accept it (it does not reach the standard of proof). But the judge cannot accept it without reasons to believe it (even though the judge may not, in fact, believe it).

Finally, the distinction between “belief” and “acceptance” can also shed light on an ambiguity in the expression “*p* being proven”. This could refer to there being reasons for believing “p” (in which case it has a descriptive meaning) or to there being reasons for accepting “p” (in which case it is once again ambiguous: it can be a description stating the existence of these reasons or it can express the performative that establishes “p” as proven).[10]

4.4. Is a precise, objective standard of proof possible?

[10] This is relevant for the discussion of the illocutionary strength of evidentiary statements. See Diego Dei Vecchi, “Acerca de la fuerza de los enunciados probatorios. El salto constitutivo”, in *Doxa*, no. 37, 2014.

The great problem raised by the standard of proof concerns finding an objective formulation for it. According to Larry Laudan's well-known critique[11], the formulas offered by our judicial systems, at least in the criminal sphere ("beyond all reasonable doubt", "sufficient incriminating evidence", "deep conviction") are vague ad imprecise. In the end they depend on the subjective considerations of the judge or jury, without guidance from rational criteria.[12]

To consider this kind of criticism in detail it seems important to distinguish two types of problems with standards of proof concerning the ambiguity of the terms "objective" and "subjective". Sometimes by "subjective" we mean the subjective attitudes or discretionary mental states of the person judging, as is the case with the standard of "deep conviction". It is enough for the judge to be convinced, without such a conviction needing to be rational in order to justify a statement that the facts are proven. This makes the criterion an arbitrary one. Other times, we use "subjective" in the sense of "vague" or "imprecise" (because, if it is imprecise, the judge ends up deciding by using his own discretion, in accordance with his own subjective criteria). Now, a concept can be affected by two types of imprecision or vagueness. We might call these intensional vagueness (the necessary and sufficient conditions for the application of the concept are not properly determined: for example, the defining notes of "book", "vehicle", etc. are not determined) and gradual vagueness (one of the defining notes of the concept is gradual – in other words it can be possessed to different degrees – as with "baldness", "tallness", "heat" or "degree of confirmation"). For example, when we speak of "reasonable doubt", although this term is given objective meaning, what is meant by "reasonable" remains to be specified, and using another gradual property must be avoided. A satisfactory standard of proof should, then, determine whether the "degree of confirmation" of a hypothesis is sufficient to be accepted. Some conditions are necessary to achieve this: (1) it must be done appealing to objective criteria rather than mental states, (2) it must be intensionally precise and (3) it must deal with the problem that "degree of confirmation" is a gradual (and unquantifiable) concept. If this issue is not resolved, this consideration cannot be used as a "threshold" or sufficiency criterion. Is all this possible? I believe it is possible to interpret the standards so that they do not depend on subjective mental states, but I am much more sceptical about the possibility that the two forms of vagueness can be satisfactorily reduced so that the judge's discretion is rendered unnecessary.[13] I will try to demonstrate the reasons for my scepticism.

---

[11] Larry Laudan, "Por qué un estándar de prueba subjetivo y ambiguo no es un estándar", Doxa. *Cuadernos de filosofía del Derecho*, no. 28, 2005.

[12] To summarise it in the words of Juan Carlos Bayón, the standard must not be subjective, or express this even covertly. Juan Carlos Bayón, "Epistemología, moral y prueba de los hechos. Hacia un enfoque no benthamiano", *Analisi e diritto*, 2008.

[13] Laudan has offered various examples of a standard of proof which would be preferable to traditional ones (for illustrative purposes only – I am not defending them here – they include

4.4.1 The “mathematical probabilism” route

An initial route consists of trying to quantify the level of credibility of the hypotheses. That means finding a method for mathematically expressing the confidence we have in them. Susan Haack has called the attempt to do this in the sphere of judicial proof “legal probabilism”. If it were objectively possible to mathematically quantify the support that the elements of judgment provide for the hypothesis to be proved (if we could say, for example, that, given certain evidence, the hypothesis is 70% or 90% confirmed, for example) then an objective standard of proof could be established, although the actual figure would have to be specified. The problem with this attempt to offer an objective standard is that it lacks satisfactory instruments for making the calculation. Attempts to apply Bayes’ Theorem to the calculation of the degree of credibility of a hypothesis (which are the most serious attempts at “legal probabilism”) seem to arouse tremendous difficulties. Bayes' theorem tries to measure the impact that a particular piece of evidence (or a set of them) has on the probability initially attributed to a hypothesis, without taking into account the evidence in question. To use it as a standard of proof, the *a priori* probability assigned to the hypothesis of guilt would first have to be established. The application of the formula would then indicate, given the impact of the new evidence, the *a posteriori* probability of this hypothesis. If we established the standard of proof at 95%, for example, we would pass the test if the *a posteriori* probability was the same or higher. So how do we assign the *a priori* probability?

---

the following:
(a) “If prosecution evidence or testimony that would be difficult to explain if the defendant was innocent is credible, and exculpatory evidence or testimony that would be very difficult to explain if the accused was guilty is not credible, then find him guilty. Otherwise acquit him.”
(b) “If the prosecution’s story concerning the crime is plausible and you cannot imagine a plausible story showing the accused to be innocent, then find her guilty. Otherwise acquit her.”
(c) “Decide whether the facts established by the prosecution refute any even marginally reasonable hypothesis you can think of concerning the defendant’s innocence. If they do, you must find him guilty. Otherwise, you must acquit.” (Larry Laudan, “Por qué un estándar de prueba subjetivo y ambiguo no es un estándar”, p. 108).

I believe it is clear that these proposals do not manage to overcome the use of subjective factors (such as the capacity to imagine stories or think of plausible hypotheses in favour of innocence). Nor do they make much progress in reducing intensional vagueness (they refer to vague notions such as “if it is credible”, “plausible”, “difficult to explain” without indicating when it must be credible, plausible, and so on). Finally, they do not solve the problem of gradual vagueness, as the concepts referred to – credibility, plausibility and so on – are clearly gradual concepts.

In some cases, statistical data can be obtained to help with this initial assignment of probability but in the vast majority of cases, the assignment of the *a priori* probability which finally determines the *a posteriori* probability is utterly subjective. Therefore, the standard of proof constructed in this way does not overcome the problem of subjectivity, it just transfers it to another point. The criticisms of authors like J. Cohen, Susan Haack, Michele Taruffo, Larry Laudan or Jordi Ferrer show the implausibility of this method of constructing an objective standard of proof.[14]

4.4.2 The problem of intensional vagueness: formulating a standard based on a selection of evidence evaluation criteria.

Let us suppose we have reached a consensus with respect to a closed list of evaluation rules (for example, the 12 already proposed). We could then require that the evidentiary argument should meet a minimum number of criteria depending on how strict we want the standard to be. However, for this to work we would need to be convinced that the conclusion of an inference complying with a greater number of these rules would always be more justified than one complying with fewer of them. And this is not the case. We have already seen that each of these rules can be complied with to different degrees, so it would be possible for hypothesis H1 to meet nine criteria, but to a lesser degree, and another hypothesis, H2, to meet just five, but to such a high degree that this would make up for the lower number of rules complied with. That could make H2 more justified than H1. Solving this requires having criteria to determine the degree necessary for *accepting* that a rule has been complied with. However, then we run up against the need for a "standard of compliance" with each rule, which would reproduce our problems (and begin a regression to the infinite).

Another possibility would be to select some of these evaluation rules as particularly relevant (or strict) and make acceptance of the hypothesis depend on compliance with these. The criterion would then no longer be quantitative (complying with a particular number of rules), it would be qualitative: rules 7, 10 and 12 must be complied with, for example. This strategy raises similar difficulties. Firstly, it is difficult to determine which criteria or rules are more important or why, because this involves deciding that compliance with the other rules cannot make up for failing to comply with those chosen. Secondly, once again a standard would have to be established to determine whether the rules have been complied with sufficiently (we must remember that different degrees of compliance are possible). Thirdly, if acceptance requires only three rules of epistemological rationality, does this mean that the others, however intuitive they appear, are irrelevant?

---

[14] See, for all these, Michele Taruffo, *La prueba de los hechos*, Ed. Trotta, Madrid, 2002 and L. J. Cohen, *The probable and the provable*.

I think Jordi Ferrer's proposed standards of proof run into these same problems. One of his formulations, for example, would be "the following conditions have to be met for the hypothesis of guilt to be considered proven:

1) The hypotheses must be proved to a high standard, explain the available data and be capable of predicting new data which has, in turn, been corroborated.

2) All other plausible hypothesis explaining the same data and compatible with innocence must be refuted."[15]

It is easy to see that Ferrer's strategy consists of formulating the standard of proof based on certain evidence evaluation criteria. In the example mentioned, these are that hypothesis makes it possible to predict new data (coinciding with our evaluation rule 8), that it has explanatory capacity (rule 10) and that it eliminates alternative hypotheses (rule 12). My argument, again, is that, *given the fact that the determination of the degree of confirmation of a hypothesis requires holistic evaluation, no subset of the evidence evaluation criteria is, on its own, a sufficient or necessary condition for achieving a certain degree of confirmation. Therefore, it is impossible to be sure that this standard requires a higher (or lower) degree of confirmation than would be required by a standard based on other evaluation criteria.* The confirmation threshold that a standard of this kind attempts to set can also be reached by a hypothesis that does not meet that standard. It therefore gives us no assurance that we are minimising the risk of error in finding guilt to a higher degree than another standard based on a subset of the evaluation criteria. In the words of Susan Haack: "As the quality of evidence has various different dimensions (...) and there is no way of ordering relative success or failure through these different factors, there is not even any assurance of a linear order of degrees of guarantee."[16] Any standard of this kind would involve setting arbitrary conditions with no guarantee that we are making proof more (or less) difficult. On the contrary, if we try to formulate the standard including all evaluation criteria it either leads to vague formulations (choosing the hypothesis that is the best or most credible explanation, for example) or prevents a distinction between the evaluation of evidence and the standard of proof.

4.4.3 The problem of gradual vagueness: refuting alternative hypotheses, an all-or-nothing principle?

It might be thought that the degree of strictness involved in refuting or eliminating the hypotheses in favour of innocence included in the third of Laudan's and Ferrer's proposals is

[15] Jordi Ferrer, "Los estándares de prueba en el proceso penal español", *Cuadernos Electrónicos de Filosofía del Derecho*, no. 15.

[16] Susan Haack, "El probabilismo jurídico. Una disensión epistemológica", page 80, in Carmen Vázquez (ed.), *Estándares de prueba y prueba científica. Ensayos de epistemología jurídica*, Marcial Pons, 2013.

a more precise standard and could operate as a “threshold”. The idea of refuting all alternative hypotheses is intriguing because, firstly, it looks like a strict criterion and, at the same time, it might be thought that refuting a hypothesis was an “all-or-nothing” matter. However, it is not that simple. Refuting a hypothesis consists of showing that there is a fact incompatible with the hypothesis (direct refutation) or with a hypothesis deriving from it (indirect refutation). The pattern of these arguments is as follows:

Direct refutation of hypothesis A:

1) We assume Hypothesis A.
2) Hypothesis B is incompatible with Hypothesis A (they cannot simultaneously be true).
3) Hypothesis B is accepted as proven (it is sufficiently confirmed).

Hypothesis A is therefore false.

Indirect refutation of hypothesis A:

1) If Hypothesis A is true, Hypothesis B will be too (Hypothesis A implies Hypothesis B).
2) Hypothesis B is incompatible with Hypothesis C (they cannot simultaneously be true).
3) Hypothesis C is accepted as proven (it is sufficiently confirmed).
4) Hypothesis B is false.

Therefore (by modus tollens) Hypothesis A is false.

As can be seen in both cases, to refute or eliminate a hypothesis a supposedly incompatible fact must be *proved,* but because of the nature of evidentiary argument we will not have absolute certainty about it. When we say that hypothesis H1 has been refuted, what we are really saying is that it is not the most probable hypothesis, because we concerned here with probabilities rather than certainties. Instead, we are saying that hypothesis H2, which assures the existence of a fact directly or indirectly incompatible with H1, seems to us more probable. So refutation is also a comparison between various hypothesis, it is also gradual and it once again requires a standard of proof. Moreover, if refuting the hypotheses compatible with innocence means showing that another hypothesis (that of guilt) is more plausible, it might then be thought that refuting the acquittal hypotheses

is no more than the other side of the coin to confirming the conviction hypothesis, rather than being a different judgment.

4.4.4 The problem of measuring the effectiveness of the standards for distributing the risk of error.

There is another problem raised by the above strategies (including the one using mathematical probability): once the standard has been chosen, how can we know the consequences of the intended result for the distribution of error? If the purpose of the standard is to try to fix a certain ratio between the number of false convictions and the number of false acquittals, we need a criterion (which must be different from the one provided by the standard) to check that the standard is producing the desired effect. But in the case of trials we do not have these criteria. Once the evidence has been evaluated and it has been established that hypothesis H is the best confirmed and that its degree of confirmation is sufficient, we no longer (except in the few cases where new more solid evidence appears and case is reopened) have another way of establishing whether or not H is true. As Bayón says, once the standard has been formulated "there would be no guarantee that the distribution of risk considered to be correct would be the exact result of its application". This is an important difference between acceptance criteria or standards from other spheres (medicine, science or even everyday life) and Law. During my life, I have learned through experience (sometimes hard experience) that if I rely on the cinema opening time information given in newspapers to find out what time the film I want to see starts, or on what the weather forecast says to find out what the weather is going to be like in a couple of days' time, I might end up missing the film or organising a disastrous picnic in the country. But I learn from reality. In Law we rarely have this chance of confirmation.

It seems to me that the above considerations point to the fact that, although it is possible to eliminate directly subjective references from the standards intended to indicate the degree of proof required, it is not possible to formulate them precisely so that the associated intensional vagueness and gradual vagueness can be overcome. The route towards a precise standard of proof appears to be blocked.

However, not all the standards we have are entirely useless or counterproductive and some are at least minimally informative. Everyone understands that the confirmation criterion "beyond all reasonable doubt" is more demanding than that of "overwhelming evidence" or "clear and convincing evidence".[17] Or that when incriminating evidence is required in gender

[17] The standard of reasonable doubt is not as useless as it might seem. Daniela Accatino has shown that the best interpretation of this standard is not the subjectivist one – there must be a conviction if the judge, in fact, has no reasonable doubt. He advocates an objective interpretation – the doubt must

violence crimes the fact that the victim “does not lack credibility” and the “corroboration with other information” of a statement establish stricter requirements than if a simple statement were enough.[18] Standards should at least provide information on whether the judicial authorities want them to be strict or not, even though they do not specify how strict, so that the idea of sufficient proof still depends on the good judgment and consideration of judges.

5. Final reflection

As we have seen, the problem of formulating the standard of proof arises basically in cases of free evaluation of evidence. Moreover, it is a problem caused by abandoning weighted evidence systems. The excessive rigidity of these systems leads to judges being delegated (1) the responsibility of evaluating the evidence and (2) the responsibility for determining whether the hypothesis is sufficient to be considered proven. The current search by evidence theorists for a more precise standard of sufficiency is, perhaps paradoxically, an attempt to maintain (1) but reduce (2). Now, if it is not possible to satisfy (2) by means of a precise standard of proof, one way of reducing discretion concerning the estimation of sufficiency consists of reducing discretion in evaluation. If rules are introduced to authoritatively determine the evidentiary result, they will simultaneously do (1) and (2). I am not suggesting here that weighted evidence and the standard of proof are the same thing, but I am alluding to the fact that one solution to reducing discretion in determining the sufficiency of the degree of confirmation consists of going back to legally or jurisprudentially weighted evidence or similar procedures. But then we once again run up against the reasons for avoiding legally weighted evidence, such as excessive rigidity. Perhaps freedom in the evaluation of evidence is simply not compatible with the precise regulation of decision-making standards and we must find other ways of distributing the costs of error.

---

be justified – which could be identified with the elimination of the acquittal hypotheses.” Daniela Accatino, “Certezas, dudas y propuestas en torno al estándar de la prueba penal”, *Revista de Derecho de la Pontificia Universidad Católica de Valparaíso*, no. 37, 2011.

[18] Along these lines, Mercedes Fernández López has proposed abandoning the attempt to specify a general standard of proof. Instead, she calls for the specification of requirements that must be met by the means of proof in each type of case so they can be considered incriminating evidence. Mercedes Fernández López, “La valoración de las pruebas personales y el estándar de la duda razonable”, *Cuadernos Electrónicos de Filosofía del Derecho*, no. 15.