ABSTRACT

This paper aims to shed light on the main problems which both forensic science and juridical science face in view of the widespread use of scientific evidence at trials. The underlying central idea is that the main challenge arisen by the use of scientific evidence in legal proceedings is not that of modelling specific standard of proof for such evidence but that of demystifying it.

The mystification of scientific evidence has been prompted by two overestimations: the epistemic overestimation and, mainly, the semantic overestimation. The first refers to the tendency to overweight the probative value granted to scientific evidence, whereas the second refers to an incorrect way of interpreting what the results of scientific proof are saying. The eradication of the latter is harder and is expressed in the abandonment of individualization paradigm and the adoption of the likelihood paradigm.

The likelihood paradigm, as opposed to the individualization paradigm, rests on the distinction of three basic questions arising when a scientific test has been carried out: 1°) What does the data say as evidence?; 2°) What is to be believed from it? y 3°) What is to be done? The first question must be answered by the expert in terms of a likelihood ratio which represents the strength or probative value of that evidence. How to convey that ratio in the expert report conclusions so that it can be properly interpreted by judges in order to assess evidence is one of the main challenges forensic science faces nowadays. The second question concerns the assessment of evidence and must be answered by judges combining the scientific evidence reported by the experts and all the rest of the available evidence. It is a challenge for both legal science and legal theory to design assessment tools and appropriate models in order to rationalize that task and guarantee, as much as possible, the accuracy of the results. The last question concerns the standards of proof, which offer criteria to act in specific circumstances or, more specifically, provide judicial guidance when a hypothesis shows a certain level of probability.

Bearing all this in mind, it could be said that a particular requirement in modelling a standard of proof when scientific evidence is provided does not seem to exist. The only specific requirements which do exist are those related to the determination of validity and reliability of scientific evidence and to the determination of its probative value.
Summary.

1. SCIENTIFIC EVIDENCE. OVERESTIMATIONS AND PARADIGMS. a) The myth. b) Epistemic overestimation. c) Semantic overestimation: from the *identification paradigm* to the *likelihood paradigm*.  

2. THE LIKELIHOOD PARADIGM. 

3. THE CHALLENGES. a) On admissibility of scientific expert testimony. b) On how to report the results of proof. c) On assessing evidence.

1. SCIENTIFIC EVIDENCE. OVERESTIMATIONS AND PARADIGMS

a. The myth

The constant scientific and technical breakthroughs have had a huge impact on the field of fact-finding and *scientific evidence* is becoming increasingly important in all judicial proceedings. These breakthroughs have been particularly important in Forensic Genetics and have marked a radical change in solving numerous legal problems, such as paternity biological research and identification of persons or criminal marks (blood stains, saliva, sperm or hair). The potential of genetic prints has reached such a level that its use in legal proceedings has become an everyday issue. However DNA proofs are not the only ones to play such a role. Voice identification, fingerprinting and ballistics, among others, are also important. Voice identification analysis, for example, has prompted convictions in terrorism cases in Spain, where the main evidence used against the accused was a telephone call stating the whereabouts of an explosive device. Ultimately, scientific evidence has become the key in proving many facts that otherwise could not have been proven.

Yet the importance of scientific proof in courtrooms has not been accompanied by a parallel process of cautious questioning; quite the contrary. Precisely by the fact that these proofs are presented as “scientific” (and because in the majority of cases, at least in Europe, come from official forensic labs), they are surrounded by an aura of infallibility which has hindered any attempt at critical reflection1, thus resulting in both their validity and probative value having

---

1 This overwhelming trust placed in the infallibility of scientific evidence has been fuelled by the tremendous impact which some television boons have had on the people, such as the American series CSI (Crime Scene Investigation).
being taken for granted. It is also clear that this conviction entails the risk that probative decisions supported by scientific evidence are considered unquestionable hence discharging the judge from making special efforts to rationally ground his decision: it is sufficient to allege the existence of scientific evidence pointing in the direction of the final decision.

The (almost) untouchable nature of this probative universe is based on two overestimations. On the one hand its results are frequently overweighted or even regarded as infallible truths, and on the other hand they are considered to say one thing when really they are saying another\(^2\). The first is an *epistemic overestimation*, the second is –so to say– a *semantic overestimation*\(^3\).

b. Epistemic overestimation.

Why is the infallibility of scientific proof so strongly accepted? Why is scientific evidence not submitted to the same critical eye as non-scientific evidence? I am not going to tackle this question in great detail. I will merely point out that the reason of such overestimation lies in the fact that scientific proof is believed to pass through epistemic and inferential basis different than those of the other proofs; more precisely, it is accepted that whereas non-scientific proof is developed through an *inductive* form of reasoning, scientific proof is developed through a *deductive* form of reasoning. What it is held, in other words, is that *non-scientific* proof is based on weakly grounded probabilistic laws, hence its results are fallible and are to be expressed in terms of mere probability. However *scientific* proof is argued to be based on universal laws which additionally are applied within a framework of scientific methodology, thus resulting in

---

\(^2\) The overestimations are denounced in the NRC Report, *Strengthening Scientific Evidence in the United States: A Path Forward*, The National Academies Press, Washington (USA), 2009, which says: “For decades, the forensic science disciplines have produced valuable evidence that has contributed to the successful prosecution and conviction of criminals as well as to the exoneration of innocent people… Those advances, however, also have revealed that, in some cases, substantive information and testimony based on faulty forensic science analyses may have contributed to wrongful convictions of innocent people. This fact has demonstrated the potential danger of giving undue weight to evidence and testimony derived from imperfect testing and analysis. Moreover, imprecise or exaggerated expert testimony has sometimes contributed to the admission of erroneous or misleading evidence” (Introduction, Paragraph: Challenges Facing the Forensic Science Community, p. 4).

\(^3\) Regarding both overestimations cfr. in more detail Marina Gascón, “Prueba científica: mitos y paradigmas”, *Anales de la Cátedra de Francisco Suárez*, Granada, (forthcoming).
unquestionable results. In short, the knowledge obtained in courtrooms is fragile whereas that obtained in forensic laboratories is quite a different.

Naturally this is a mistaken conviction. Scientific proof in general does not constitute deductive reasoning but is based on mainly probabilistic laws. Furthermore, its results must be interpreted using statistics, hence it would be difficult to speak of “objectivity”\textsuperscript{4} let alone infallibility regarding conclusions made from them.

It could be the case that the myth of infallibility of scientific evidence is linked to the great success of DNA proof, whose methodology is so developed and its accuracy so high that in practice it can be treated “as if” it were infallible. However, even if that were true, a more critical reflective attitude should be adopted for two reasons.

The first is related to the fact that epistemic quality of the results of scientific proof depends on various factors. Firstly, it depends on \textit{scientific and/or methodological validity}: many of these proofs can be carried out using different scientific methods, not all of which are considered to be equally reliable. Secondly, the reliability attributed to scientific evidence also depends on \textit{technical quality}. It is worth making a distinction here. On the one hand, we can speak of \textit{technical-procedural correction}, referring to the whole process leading from the discovery of the mark, sample or element to be analyzed until its final analysis in the lab (for example, in order to attribute reliability to the result of a fingerprint proof the problem lies not only in the scientific validity of the proof but firstly knowing who took the fingerprint, under whose orders, where exactly it was kept, the details of the chain of custody, etc. The same applies to a blood

\textsuperscript{4} “Any probability judgement in a particular case, even when the judgement is frequency based, has a component based on personal knowledge”. That is the same as saying that any probability judgement is essentially personal and hence subjective. Understanding statistical data as objective data, in the sense that they are incontrovertible and universally achievable, shows an unreal conception of the possibilities of science. It is possible to speak of objectivity as an inter-subjective agreement. In this sense it is easier for scientists to accept a result if it is based upon relevant statistical data than if it is based upon subjective assessments of probabilities, but this does not amount to saying that scientists believe that result to be incontrovertible. (F. Taroni, C. Aitken, P. Garbolino and A. Biedermann, \textit{Bayesian Networks and Probabilistic Inference in Forensic Science}, John Wiley and Sons Inc., Chichester, 2006, cap. 1: “The Logic of Uncertainty”, p.21).
stain, urine or saliva sample.) On the other hand, we can also speak of technical-scientific 
correction, referring to its correct handling in the lab: very briefly, by skilled staff following 
adequate protocols. The rule here should be: “the higher are the expectations for probative 
value placed upon the proof, the more rigorous should be the controls for its performance”. 
Lastly, we must remember that proofs are carried out by human beings, and human error can 
occur. It is worth noticing in this respect that over the last few years many studies have 
highlighted cognitive risks in some scientific proofs, above all in those traditional ones, such as 
fingerprinting and handwriting, which have a strong comparative component and depend 
greatly on expert judgment.

In conclusion, validity of scientific proofs (and reliability of the obtained data) depend on 
the scientific validity of the used method, the appropriateness of the used technology and the 
adoption of rigorous standards of quality, and must not be taken for granted. Therefore even in 
relation to DNA evidence, which has been thoroughly tested and improved over time and 
appears to be the model for the rest of expertise forensic fields, there can be no doubts about 
the need to pay attention to all these issues when evaluating its reliability. The problem of 
scientific evidence, however, does not lie in the DNA test, but in the existence of many other 
fields of forensic expertise which have not received the same attention despite forming the 
grounds of many judicial decisions.

---

5 It is clear, for example, regarding a DNA test, that the state in which the samples arrive to the laboratoies is crucial: if the samples have not been appropriately collected and conserved (for instance, because they have been contaminated with foreign DNA) the efficiency and reliability of the test diminishes. Therefore, the collection of the samples must be carried out carefully, and the maintenance of the chain of custody is central in order to avoid the samples losing its probative value. It is striking in this regard the enthusiastic way in which some countries have incorporated DNA proofs without previously establishing a normative framework to regulate the procedures for collecting and conserving the samples and, therefore, guaranteeing the reliability of the results.

6 Regarding cognitive risks, let us remember the Mayfield case, in which three FBI experts and an independent expert link a fingerprint found in the Madrid train bombing inquiry to the suspect Brandon Mayfield. Soon after the link was shown to be wrong. The fact that the four experts made the same mistake can be explained because when the three FBI experts carried out the tests they already knew the independent expert’s outcome, that is, this can be explained by a cognitive bias.
Negative consequences derive from the lack of serious reflection on the epistemic statute of scientific proofs. Perhaps the most blatant and adverse of them is the absence of scrutiny in their validity and reliability, thus permitting the entry of real junk science into the proceeding, which is sometimes used by expert witnesses and laboratories as a lucrative business. Let us consider, for instance, handwriting or polygraph analysis. In addition the myth of infallibility of scientific evidence entails the risk of unawareness of miscarriages of justice which may be committed as a result of it. Finally this myth entails an additional risk, an adverse effect from a juridical point of view: since the forensic report is assumed to be infallible, a judge cannot oppose it without strong reasons, thus being possible to say that it is the expert himself who tells the Judge what to believe regarding the hypothesis under consideration. In that way, however, the role of fact-finder has shifted from the judges of the trial to the expert, and therefore a new system of proof grounded on the experts’ authority will have been adopted.

c. Semantic overestimation

Not only does the infallibility dogma of scientific evidence have to be reviewed but also what I previously named the “semantic overestimation”, that is the fact of considering results of the proof (not only as being infallible but also) as saying something other than what they really are saying. Let us consider this in more detail.

If scientific evidence has been idolized in legal proceedings it is because it has been taken for granted that its results directly speak about what is to be proven. It is thought, particularly, that the result of a DNA proof shows directly whether the analyzed sample derives from (or does not derive from) the same source as the compared sample; that the result of a ballistic proof directly shows whether the bullet came from (or did not come from) the defendant’s gun; that the shoe-mark comparison test tell us whether the examined shoeprint comes from (or does not come from) a certain person; that the result of a handwriting test shows whether the

---

7 The Innocence Project, set up by the attorneys Barry Scheck and Peter Neufeld in the Cardozo Law School in order to demonstrate the innocence of a great deal of condemned people by DNA proofs, has highlighted not only the weaknesses of the traditional means of proof, such as testimonies and confessions, but also the weaknesses the scientific proofs supporting some of the convictions. See on www.innocenceproject.org
examined handwriting sample belongs (or does not belong) to the accused; that a voice comparison test tell us whether the voice of the caller alerting as to the whereabouts of the bomb belongs to (or does not belong to) the accused, etc. In short it is thought that the result of a scientific test speaks in the same terms in which the Judge must speak.

This belief represents what in forensic field is called individualization paradigm, which appears to be the pattern to follow for all the criminal identification techniques. Such a paradigm is settled in the belief that the result of a scientific test can link the analyzed mark to a specific source, excluding all others.

However things are not so. In fact the individualization paradigm has come under harsh criticism from many scientists and scholars for decades, who have held that although it may be a simple and intuitive pattern of analysis, it is not acceptable: the unproved claim that a forensic practitioner can link an unknown mark to a unique source represents a wrong probabilistic intuition equating infrequency with uniqueness. In short: obtaining definitive proof of uniqueness is considered impossible. Therefore “a dispassionate scientist or judge

8 Referring to the identification paradigm, the NRC Report Strengthening Forensic Science in the United States says that analysts use to believe that “Some forensic science methods have as their goal the “individualization” of specific types of evidence (typically shoe and tire impressions, dermal ridge prints, toolmarks and firearms, and handwriting). Analysts using such methods believe that unique markings are acquired by a source item in random fashion and that such uniqueness is faithfully transmitted from the source item to the evidence item being examined (or in the case of handwriting, that individuals acquire habits that result in unique handwriting). When the evidence and putative source items are compared, a conclusion of individualization implies that the evidence originated from that source, to the exclusion of all other possible sources” (Chapter 1, Introduction, Sección Pressures on the Forensic Science System, Paragraph: Questionable or Questioned Science, p. 43).

9 One of the most stimulating essays regarding the use of statistics in forensic science, the work of R. Royall, Statistical Evidence: A likelihood paradigm (Monographs on Statistics and Applied Probability, Chapman&Hall / CRC, London, 1997, Preface, p. xi) explicitly denounces that the use of statistic methods often leads to misinterpret the outcome of the tests not because the experts use the statistics wrongly but due to the use of the individualization paradigm.

reviewing the current state of the traditional forensic sciences would likely regard their claims as plausible, underresearched and oversold”. 11

As a result scientists have called for the need to interpret the results of the tests, not in identification terms (of a mark to a source) but in terms of likelihood. It is argued more specifically that there is no scientific grounding for an expert to hold categorically he has been able to identify a person or an object from the analysis carried out in the laboratory. This criticism, however, goes even further: there is neither scientific grounding that the expert report carries out a probabilistic assessment regarding the possibility of linking the examined mark to a person or an object. The outcome of a voice analysis, for example, or a comparison of DNA profiles, does not demonstrate (neither conclusively, nor probably) that the analysed voice or DNA sample belongs to one person or another but only gives data that, once interpreted with suitable statistic tools, say things such as the following: <<it is X times more probable that this feature appears if the voice belongs to the accused than if it does not belong to him>>; or <<it is X times more probable that DNA profiles match if the analysed mark belongs to the accused than if it does not belong to him>>. In other words, scientific evidence speaks about probability of analytical and technical data resulting from lab tests in face of the considered hypothesis, and not the opposite case; that is to say, they do not speak about the probability of the considered hypothesis in face of the data. Therefore the widespread custom of expressing the outcome of scientific tests in terms of categorical or even probabilistic identification of the hypothesis under consideration (particularly widespread in traditional forensic techniques) must be abandoned. Besides, as Ch. Champod highlighted in his harsh criticism of individualization paradigm, forensic science, in order to be useful, does not need to deal with this paradigm 12.


2. THE LIKELIHOOD PARADIGM

Yet a question arises then: If the individualization paradigm turns out not to be valid, how should the forensic experts assess evidence? The likelihood paradigm named by R. Royall has been put forward to answer this question\textsuperscript{13}. Such a paradigm rests on the need to clearly distinguish between the three following questions that can be put once the pertinent analysis and observations of a scientific test have been carried out and some data has been obtained: 1) what does that data say about hypothesis A versus hypothesis B; or how should these observations be interpreted as evidence regarding A versus B?; 2) what is to be believed from here? And 3) what is to be done?\textsuperscript{14}.

That distinction constitutes a remarkable tool to re-construct the interaction between the judicial process and a kind of proof which is developed outside it. In fact, this paradigm allows us to draw a net delimitation of tasks between judges and experts.

\textit{What the data says as evidence?}

Interpretation of data as evidence is the main task of the experts in the trial, which involves interpreting what the outcome of the test says with respect to a hypothesis as opposed to the contrary. In this way, for example, in a voice comparison proof the expert ought to interpret the outcome of the test (let us suppose positive: the compared voices match) valuing it in relation to both hypotheses (let us suppose that A: “the voice belongs to the accused”; and B: “the voice does not belong to the accused”). Let us recall that this interpretation ought to be done in terms of likelihood ratio, and it is of utmost importance that it be done correctly, since it is what must be reflected in the conclusions of the expert report and therefore what is communicated to the judge.

\textit{What is to be believed?}

\textsuperscript{13} R. Royall, \textit{Statistical Evidence: A likelihood paradigm}, cit.

\textsuperscript{14} A more detailed exposition of this paradigm in “Razones científico jurídicas para valorar la prueba científica: una argumentación multidisciplinar” (M. Gascón, J.J. Lucena y J. González), en \textit{La Ley}, núm. 7481, 04.10.2010.
Once the data has been interpreted and suitably expressed in the expert report a second question must be answered: what is to be believed from the data? That is to say, the reliability of the opposing hypothesis (A and B) must be assessed and that must be done not only considering what the data says but also what the other available evidence says.

The assessment of the hypothesis is undoubtedly granted to the judge for at least two reasons; first of all because it is the judge who has the institutional role to provide a legal decision in a controversial case, and hence the role of previously state the relevant facts of the case. Secondly because the free conviction principle requires the judge not to accept a fact until he is entirely convinced of it, and as a consequence it requires the judge (not the expert) to determine what is to be believed about the hypotheses under consideration from the evidence provided, including scientific evidence. In this regard it should be noted that the individualization paradigm is not compatible with exclusively assigning the assessment function to judges. In effect, if the expert report spoke in terms of <<mark m comes (or does not come) from source s>> it would have been the expert (not the judge) who would have said what to believe about the hypotheses.

In short, the net distinction between the expert’s task (interpret and communicate what the data says) and the judge’s task (assessing the data in the light of the other available evidence) is the core of the likelihood paradigm. The individualization paradigm, on the contrary, does not distinguish them since it prompts an expert report in terms of <<mark m comes from source s (for example the accused)>>, which are the terms that must be used by the judge when assessing evidence. In that way, however, important legal principles (such as the free conviction principle) are not observed.

What is to be done?

15 Otherwise, this paradigm entails an important risk of fallacies, such as the well known prosecutor’s fallacy. About the prosecutor’s fallacy and the defendant’s fallacy, cfr. Thompson, W. y Schumann E., “Interpretation of statistical evidence in criminal trials. The prosecutor’s fallacy and the defense attorney fallacy”, Law Human Behav. 1987, 11, pp.167 ss. About the prosecutor’s fallacy it can be seen more recently Leung, W. C., “The prosecutor’s fallacy – a pitfall in interpreting probabilities in forensic evidence”, Medicine Science and the Law, 42, 2002, 44-50.
Lastly the likelihood paradigm also points out the distinction between *what is to be believed* and *what is to be done* after a scientific test has been carried out. While what is to be rightly believed about a hypothesis is expressed as the probability of that hypothesis and *coincides* with the *assessment* of evidence, what is to be *done* (that is, making decisions in light of that probability) states the issue of the *standards of proof*, that is the issue of determining the degree of probability which a hypothesis must reach in order to be considered proven and act accordingly. It should be noted in this regard that, while the assessment of evidence is an issue which completely relies on epistemic rationality, setting standards of proof is a matter of *policy* (resting on values or political options), entirely depending on how tolerant the system is willing to be with the two possible errors which can be made when making a decision: accepting as proven what is false and accepting as not-proven what is true. It is clear, for example, that the standard of proof in civil proceedings is lower than in criminal proceedings. Therefore while in the first case it could be accepted that the proof of a fact lies upon preponderant probability (let us say 55%) in the second case we are only willing to convict upon a highly qualified probability. This is due to the reason that we are less tolerant with the error of convicting an innocent person than with the error of absolving a guilty person.

To sum up, *what is to be believed* about a hypothesis depends on what the data says about it as well as the rest of available evidence. *What is to be done* obviously depends on what is to be believed but also on the normative framework in which the decision is to be adopted. Therefore, with the same belief it is not the same to make a decision in civil proceedings than in criminal proceedings. Furthermore, in the last case it is not the same to make a decision in legal inquiry proceedings than at the trial.

### 3. THE CHALLENGES

Important challenges for accurate decision making when scientific evidence is used derive from the likelihood paradigm. One of these concerns the assurance of the scientific validity and reliability (in short, the quality) of evidence admitted at trial. Another is related to the content
and sense of the expert witnesses report. The last one deals with the way of assessing evidence when scientific data is provided.

a. On admissibility of scientific expert testimony. The judge as gatekeeper?

There is no doubt about the epistemic interest of assuring that expert testimony admitted at trial is grounded in scientifically reliable methodology, since that control aims to keep pseudoscience out of the decision making field. How to achieve this goal is what differentiates legal systems.

In this respect two main positions may be adopted. On the one hand, it is possible to hold that no specific control of scientific admissibility is needed along with the control of procedural admissibility (namely, relevance and legality), so that all the relevant and legally gathered evidence ought to be admitted at the trial, postponing the question of its scientific validity and reliability until the moment of assessment. On the other hand it is also possible to hold that there must be a control of scientific admissibility along with the control of procedural admissibility and different from the phase of assessing evidence, that is of attributing it a probative value.

European countries in general have chosen the first position. Relevant and legally obtained scientific evidence enters the trial and is to be assessed by judges, who confer it greater or lesser probative value depending, among others, on the reliability of the methodology used to obtain it. Yet this option has some pitfalls.

Firstly, we should mention a central conceptual issue which nevertheless has practical effects. This model does not distinguish between the phase of admitting expert evidence and the phase of attributing probative value to it, thus resulting that the scientific reliability of the evidence.

16 Ultimately, what is held is that if the evidence is not scientifically reliable, it will be deprived of any probative value when assessed by the judge.

proof is often confused with its **probative value**. It should be noticed, however, that admitting expert evidence does not presuppose ruling on its probative value but purely on the possibility of being taken into consideration in order to decide the case. Both concepts (scientific reliability and probative value) must be clearly distinguished.

Secondly, there is also the problem that judges usually lack the required scientific education to evaluate the validity and reliability of the proofs. Therefore leaving this task entirely to the judges free assessment, without establishing any objective standard to give them guidance, may well lead to either dismiss scientifically valid evidence, which otherwise might have well contributed to discovering the truth, or –what is more probable- to give credit to some proofs which are merely pseudoscientific charlatanism and can prompt erroneous judicial decisions. Scientific evidence has great persuasive power on judicial conviction, hence if that evidence entered the trial without any control it could risk making a decision based on mere junk science\(^\text{18}\). That is why it seems to be more appropriate the alternative option, namely the one that claims to establish an **objective standard** of scientific admissibility of the proofs along with the examination of the procedural admissibility.

This option has been adopted in the United States. What is known as the **Dauber test**\(^\text{19}\) has settled there as standard of admissibility, together with the **general acceptance** of the used theory and technique from the scientific community concerned, the **scientific validity** of the applied technique, which must be proved in accordance with several factors: a) if the theory or technique is refutable and testable and has been tested; b) if they have been subjected to peer review and publication; c) if the potential error rate is known; d) if standards and quality controls concerning its operation exist. These factors do not constitute a checklist but do

---


certainly represent a call for judges to carefully examine a great deal of expertise that had not
been questioned up until now, in that way leading the judges to recuperate their *gatekeeper*
role. The Dauber ruling, in short, involves an exhortation of judges to examine more critically
scientific expert testimony\(^\text{20}\). In fact, after *Dauber* traditional fields of expertise, such as
lophoscopy, handwriting examination and ballistics have started to be seriously questioned.

However the *Dauber standard* has not ended the controversy, given that its application
assumes that judges are able to understand and apply the scientific concepts which are needed
in order to examine the validity of proofs, which is not always true. Of course when a proof has
already been subjected to intense debate and accepted by the scientific community, such as
the case of DNA proof, no question arises regarding its admissibility, but there will always be
other forensic expertise scientifically weaker than DNA tests. It is not surprising, therefore, that
some have called for the need to provide the judges science education in the different fields of
expertise, especially in the forthcoming areas. In any case many questions remain open: Is it
realistic to urge the judges to acquire scientific training in order to determine the reliability of
scientific evidence?\(^\text{21}\) Given the fact that not every judge would have the same expertise, would
this not accentuate the disparity of practice? Is it not, however, more dangerous to leave that
task entirely to the experts, who could have underlying interests in promoting certain
techniques? Would it not be possible a system where judges decided with the help of an
independent, scientifically competent body?\(^\text{22}\).

\(^\text{20}\) Eloquently Ch. Champod y J. Vuille express this idea: “The Daubert, Joiner and Kumho trilogy of decisions
implicitly emphasises the scepticism which the judge must maintain vis-à-vis the expert, who is thus no longer
considered a member of an authoritative elite but a social agent comparable to any other, possibly subject to pressure
of a political and economic kind that may impair his discernment”, *Scientific Evidence in Europe. Admissibility, Appraisal
and Equality of Arms*, Comparative study on scientific evidence drawn up for the Bureau of the Council of
Europe’s European Committee on Crime Problems (CDPC) (2010), 10 (See

\(^\text{21}\) This is the reason why the *The Law Commission* proposes the possibility that in exceptional circumstances the judge
can ask for expert help in order to assess scientific validity of the proposed technique. (I refer to the report *The Admissibility

\(^\text{22}\) “One possibility might be to establish a European body on the lines of the British Forensic Science Advisory
Council and Forensic Science Regulator to act as the main adviser to political and judicial authorities on the
reliability of the scientific techniques that are used (...). The more a particular form of scientific evidence is deemed,
of itself, to be critical, the more necessary it becomes to subject it to the rigorous assessment of an independent
b. On how to report the results of scientific proof

Besides the central issue of the reliability of scientific evidence there is another important question concerning the results of proofs to which unfortunately little attention has been paid\(^{23}\). The question concerns, more precisely, the way in which the results of the test should be interpreted and expressed by forensic experts in order to avoid them being misunderstood or overweighted by judges. In short, how the expert witnesses report should be drawn up in order to properly communicate the results of the proofs to the courts.

In accordance with the likelihood paradigm the expert should not give an opinion about the hypothesis under consideration but he should only give a rigorously scientific account of the data, thus allowing the judge to understand its exact meaning in order to assess it along with the rest of the available evidence. In this regard, a central issue should be emphasised: the conclusions of the expert report ought to determine the meaning (or the level of probability) of the analytical and technical data resulting from the test in the light of all the hypothesis under consideration, and not the opposite case, that is the meaning (or the level of probability) of the hypothesis under consideration in the light of the analytical and technical data resulting from the test. Understanding that the results of the test speak in the last terms, that is that the proofs tell the judge precisely what he has to identify in order to make his decision, is inherent to the individualization paradigm, which does not differentiate what the data says (an expert body. Such a body would not have binding powers, but could nevertheless issue recommendations to domestic courts that would assist judges when they had to rule on the admissibility of a new forensic technique or the reliability of a new form of evidence, or when it seemed appropriate to abandon a form of evidence that had become obsolete”. (Ch. Champod y J. Vuille : *Scientific Evidence in Europe. Admissibility, Appraisal and Equality of Arms*, cit., p. 109).

\(^{23}\) That was regretted not long ago by S. M. Willis, “Forensic science, ethics and criminal justice”, J. Fraser & R. Williams (Eds.), *Handbook of forensic science*, Cullompton, UK: Willan Publishing, 2009, pp. 523-545. The NRC Report *Strengthening Forensic Science in the United States* also draws attention to that point, by establishing that it is absolutely necessary “to provide the essential building blocks for the proper assessment and communication of forensic findings”: (Reporting Results Section, p. 186).
task) and what is to be believed from it (a judicial task), and leads to multiple misinterpretations and sometimes miscarriages of justice.

Probably there is no magic formula to report the conclusions of the tests and probably there is not only one formula. Anyhow it is clear that the pronouncements (categorical or even merely probabilistic) on the hypothesis under consideration should be avoided, and on the contrary pronouncements such as the following should be adopted: «It is $R$ times more probable that DNA profiles match if the analysed mark $m$ comes from the source $s$ than if it does not come from it»; or «It is $R$ times more probable that the compared fingerprints match if the doubted fingerprint belongs to the accused than if it does not belong to him».

Therefore, the so called 

Verbal Probability Scales, widely used in fingerprinting, handwriting examination, ballistics, facial comparisons, voice identification by linguistics analysis, comparisons of shoeprints, tire tracks and toolmarks are wrong. In fact these scales reproduce the identification paradigm by stating not what the data says but what it is to be believed from it regarding the hypothesis under consideration. In this respect forensic science faces an important challenge, that of adopting a standard of conclusions to give guidance to the labs in their work and at the same time to guarantee the use of common criteria\textsuperscript{24}.

Finally, let us make one more observation to that regard. Statistics plays an essential role in the interpretation of the data resulting from many scientific proofs. In fact, in most cases the data stated in the conclusions of the expert report that in turn are to be assessed by judges are statistical data. However, it is clear that people in general (and judges in particular) attribute excessive weight to statistical data (overweighting thesis)\textsuperscript{25}. Due to this fact, it does not seem out of the question to consider the need for judges to be instructed in this matter. Of course this does not mean that judges should become mathematicians or amateur scientists, but that

\textsuperscript{24} Concerning this point it is worthy of mention that the Swedish National Laboratory of Forensic Science (SKL) announced in 2009, after years of internal debate, its decision to unify the way conclusions must be drawn in order to provide the same kind of information to Courts whatever the type of evidence involved.

\textsuperscript{25} Limiting ourselves to Europe, the cases of Sally Clark, in England, or that of Lucia de Berk, in the Netherlands, absolved in 2010, are famous. Both and some others are cited as examples by Ch. Champod y J. Vuille, \textit{Scientific Evidence in Europe}, cit, p. 10.
they should be given the required skills in order to avoid the risk of misunderstanding or overestimating the statistical data which are shown in the outcome of a scientific test.

In any case, even if such instruction were beneficial, this would not diminish the important secondary role of the experts in the task of interpreting report data. To identify what the data is saying exactly and what degree of probative strength should be granted to it will always be a difficult issue. No matter how clearly and rigorously written the report is, it may be unavoidably misunderstood. That is why the experts should be in the courtroom in order to help the judges interpret the report in a correct way. In other words, the presence of experts in courtrooms is an important epistemological guarantee.

c. On assessing evidence

Assessing evidence is the core of probative decisions and consists of determining what is to be believed about the hypothesis under consideration. Therefore, assessing scientific evidence consists of attributing probative value to it with respect to the hypothesis describing the facts of the case. Two questions arise concerning the assessment: who is to assess and how to carry out the assessment.

Who is to assess?
The answer to this question may well appear as evident, and it is indeed: if it is the judge who has to provide a decision for the case, then it is also the judge who has to accomplish the previous task of stating the facts of the case. It is worth, however, insisting on that point, inasmuch as, precisely by applying the individualization paradigm, many courts are acting “as if” it were the expert witness and not the judge who assesses evidence. In effect, by expressing the conclusions of the tests in accordance with the individualization paradigm (that is by reporting in terms of, for example, “in the light of the data the hypothesis is probable in a degree x”, or “the data resulting from the scientific test makes the hypothesis probable in a degree x”, or “are compatible with the hypothesis in a degree x”, or more directly, “the mark m comes from the source s”), it is the expert (not the judge) who has stated what is to be believed about the
hypothesis. According to the likelihood paradigm, however, it is the judge who has to assess evidence, including scientific evidence.

Therefore it is worth noting that it makes no sense to say that “the judge is peritus peritorum”. According to the likelihood paradigm what judge must not do is abide to the expert report (in that way breaking the free conviction principle) or, on the contrary, abandon it in his quality of peritus peritorum (in that way assuming he has expertise which probably does not have). What the judge must do is another: state what is to be believed on the hypothesis under consideration in the light of the expert report and the other available evidence.

How to carry out the assessment?

The second and central question deals with how to carry out the assessment (that is what is to be believed about the hypothesis under consideration) when scientific evidence has been provided.

When there is only scientific evidence supporting the hypothesis, what is to be believed about it is entirely determined by (or coincident with) the probative value attributed to the scientific evidence. When, on the contrary, there is non-scientific evidence along with the scientific evidence, which is the common situation, determining what is to be believed about the hypothesis requires combining the probative value of both kinds of evidence. Naturally, this task is commonly accomplished by judges without any help from mathematics, “subjectively” so to say. However, given the fact that the probative value of scientific evidence can be expressed in statistical terms, mathematical tools apt to carry out the assessment have been proposed over the last decades.

The main attempt has been made by applying the Bayes’ Theorem\textsuperscript{26} to juridical inferences based on subjective probabilities\textsuperscript{27}. The Bayesian formula in its simplest formulation in the field

\textsuperscript{26} The Bayes’ Theorem is not the only mathematical tool which has been proposed as model for assessment. Bayesian Networks also represent a powerful tool for the probabilistic inferences in forensic science when there is multiple evidence, scientific or not. Cfr. F. Taroni, C. Aitken, P. Garbolino, A. Biedermann, Bayesian Networks and Probabilistic Inference in Forensic Science, Wiley, 2006.
of scientific evidence assessment affirms that the probability of an hypothesis H on the basis of evidence E can be determined in terms of the statistical frequency in which given H occurs E and the probability attributed to H before the new evidence has been incorporated.

\[ P(H/E) = \frac{P(E/H)}{P(E/no-H)} \]

- \( P(H/E) \): probability of H given an evidence E (posterior probability).
- \( P(E/H) \): degree of confidence that E should occur if H is true.
- \( P(E/no-H) \): degree of confidence that E should occur regardless of whether H is true.

\[ \frac{P(E/H)}{P(E/no-H)} \] (Likelihood ratio of E respect to H, which measures the probability with which E occurs being H true than being H not true).

- \( P(H) \): degree of belief that H is true before evidence E became available (prior probability assessments).

The usefulness of the bayesian formula lies on the fact that it permits combining statistical data about a certain event (that coming from the expert report and expressed in a likelihood ratio) with non-statistical data (the prior probability), and the result of the combination is the posterior probability, which expresses the result of the assessment of the evidence. In reality the Bayes’ Theorem represents the impact that the inclusion of new statistical data (such as that resulting from scientific tests) has on the previous belief in the hypothesis H that we intend to prove.

However the attempts of using the Bayes’ Theorem to improve the accuracy of fact finding are usually viewed as frustrating, since many flaws and problems have been identified. Some of the criticisms are due to practical considerations. For example, they highlight the difficulty facing the judge to quantify its subjective previous assessment (the prior probability);

---

24 “Subjective probability” indicates the degree of personal belief of a rational factfinder regarding the occurrence of an uncertain event.

or the risk which entails putting statistical tools in the hands of judges which most of times are incomprehensible to them. Furthermore, Bayesian formula entails a relatively simple mathematical calculation in the basic case of having to assess only one proof which directly speaks about the hypothesis under consideration, which in turn is a simple hypothesis. However the difficulty of the calculation increases remarkably when used to solve more complex situations, such as the plurality of evidences relating to a hypothesis, the cascaded inference or the proof of a hypothesis referring to a complex fact. Other criticisms, however, question the epistemic validity of the model in the judicial field, as it leads to contraintuitive results and would doubtfully be compatible with fundamental juridical principles, such as the presumption of innocence.

There is no doubt that these criticisms ought to be considered seriously. In any case, the pitfalls of bayesianism as a general model for assessing evidence does not impede it from being useful as a specific model for assessing scientific evidence; more precisely, as a way to quantify the probative value of scientific evidence: the Theorem is useful to clearly determine to which extend the result of a scientific test (by the likelihood ratio) contributes (raising or diminishing the prior probability) to form the judicial conviction about the hypothesis to be proven. This means that two distinct concepts of probability can coexist in the assessment of evidence. On the one hand, the mathematical or statistical probability, which is used to quantify the probative value of scientific evidence and is expressed through the likelihood ratio. On the other hand the logical or inductive probability, which can be used to express the conviction

29 The main criticisms are directed against the role played by the prior probabilities in the calculation of the final probability of the hypothesis we intend to prove. If the value of the prior probability were very low, would be also very low the value of final assessment (the posterior probability), regardless of whether the probability value indicated by the new evidence E were high, which is clearly contraintuitive. Furthermore, connected to that, if a very high value were wrongly attributed to the the prior probability, it could risk fundamental legal principles, such as the presumption of innocence. In short, the use of Bayes’s Theorem produces a overestimation of the prior probabilities and a underestimation of the new evidence in the calculation of the final probability. On the problems of bayesianism cfr. L.H.Tribe, "Trial by Mathematics", Harvard Law Rev., 84, 1971; M.Taruffo, La prueba de los hechos, Madrid, Trotta, 2004; y H.L.Cohen, The Probable and the Provable, Oxford, Clarendon Press, 1977. Me remito también, en todo caso, a la atinada y detenida crítica de L.Laudan al bayesianismo, en su trabajo de este mismo volumen.
regarding the hypothesis under consideration in the light of all the available evidence (scientific or not). A standard of proof may well be constructed upon the latter.