BURDEN OF PROOF. TRUTH OR LAW?

BY

HANNU TAPANI KLAMI
MINNA HATAKKA
JOHANNA SORVETTULA

CONTENTS

1. SOME WORDS ABOUT THE PAST

When the assertions of the parties to a lawsuit diverge, the court has to decide which party should be given credit. It is not exclusively a matter of truth: more often than not there is a genuine diversity of opinions, both parties having a truth of their own. This is particularly so when human intentions and/or their interpretation by the parties are at stake. In such situations the judge may not ask: What has been the truth? but rather: Which of the parties is right?, i.e. whose "truth" is better justified and also preferable—although not necessarily more truthful.

In such situations—and even more generally—resorting to *rules* is, in primitive law and even later, often envisaged as a method for justifying evidential reasoning, because, so it is thought, it indicates an impression of discipline and objectivity in situations where the parties disagree strongly. Resorting to rules tends to legitimate the dispute-solving function of the judiciary.[1]

* The present essay is a part of a series of papers on the theory of evidence. We here deal with the relationship between evidentiary value and sufficient evidence (evidentiary requirement). The first paper was published in *Rechtstheorie* 3/1988 (Hannu Tapani Klami, Marja Rahikainen and Johanna Sorvettula, "On the Rationality of Evidentiary Reasoning—A Model"), but we have now further developed the method presented.

Other essays in this series—which together with the empirical results of our Swedish-Finnish project group "Law and Truth" will eventually appear as a book—will be published in different domestic or foreign reviews of Festschriften:

"Legal Language and Evidence" (*Festschrift K. Opalek*, 1989)
"Evidence and Legal Reasoning" (to be published in *Law and Philosophy*)
"Evidentiary Value" (*FJFT* 1989)
"Studies on the Theory of Evidence", containing:
"Combined Mixed and Pure Evidence"
"Elimination Evidence"
"Causation Evidence" (*Oikeustiede-Jurisprudentia* 1989)
"Original Probability and Evidentiary Value" (to be published in A. Aarnio and C. Varga (eds.), *A Finnish-Hungarian Symposion on the Rule of Law*, 1990).

Believing that scientific progress results from continuous discourse, we are publishing our ideas first as separate essays—even though this means omitting or at most only mentioning many problems dealt with in our other papers. But we seek comments and criticism, in order to improve our conception, rather than imagining that the fortcoming book will at once solve all problems in a conclusive manner.

[1] On the history of evidence in antiquity, see, e.g., the papers of Henrion and Perelman in Perelman & Foriers (eds.), *La preuve en droit*, Brussels 1982.

Although the idea of *onus probandi* is old, it was only during the 19th century that one began to distinguish dogmatically between rules about the *evaluation of evidence*, *presumptions*—refutable or irrefutable—and rules about the burden of proof.[2] As early as the Middle Ages there were Roman-law-based legal maxims such as "ei incumbit probatio qui dicit, non ei qui negat" and "actori incumbit probatio—exceptione reus fit actor". But although learned law played a considerable role as far as substantive norms were concerned, it is rather difficult to assess the impact of such theories on procedural law under circumstances where oaths or access to oath played an important part, along with rather "superstitious" methods for assessing the "truth". If there is a certain line of development in the law of evidence, one may describe it as follows:

Legal rules	LRE+	LRE+P+
about evidence ↦	presumptions ↦	burden-of-proof rules
(LRE)	(P)	(BOP)

But then the legal rules about evidence (LRE) and presumptions (P) are amalgamated to a *free evaluation of evidence*, with rudiments from the earlier period. What is then left?

One may perhaps summarize the position in the following way.

(1) *Free evaluation of evidence* concerning the legally relevant facts, i.e. the *evidentiary theme*, *ET*, or its negation, $\sim ET$, respectively.
(2) Certain more or less general criteria for the establishment of ET or $\sim$ET, i.e. *sufficient evidence*.
(3) If the evidence in either direction is not sufficient, one has to apply *rules about the burden of proof*: these say what the factual premise(s) of a decision will be if there is not sufficient certainty about the facts that the parties have tried to prove.

In this paper we will attempt to analyse the conception, design and function of rules about burden of proof. Our starting-point is the idea that the evaluation of evidence, the criteria of sufficient evidence and the rules about the burden of proof are parts of a functional unity, a process that we call *evidentiary reasoning*. We intend to show how this unity can be made operational in such a manner that one need not resort to intuition or "holistic appraisal".

[2] On the history of the discussion about the burden of proof, see, e.g., the classical monograph of Leo Rosenberg, *Die Beweislast*, Munich 1965.

2. CONSTRUCTION OF BURDEN-OF-PROOF RULES AND CONCEPTUAL CRITERIA

When the principle of the burden of proof was established, that is, when it became incumbent on the party who positively referred to a fact to prove it, the main idea was simple enough.[3] "Cuius commodum, eius periculum": the party who benefitted from proving a fact had both the duty to prove it and the risk of its remaining without sufficient evidence. The principle of the impossibility of a negative proof may have been based on primitive thinking, combined with undeveloped methods for presenting evidence. It was said that only the Devil could prove a negative fact; such a "probatio diabolica" was perhaps excluded also because according to Mosaic Law two reliable witnesses constituted full proof. But it is difficult to think that two witnesses could have been used to prove e.g. that the defendant had *not* committed the crime. They could only testify that they had not seen the defendant commit the crime. When the exclusion of guilt by means of an alibi was only dimly understood, one preferred to operate with requirements about positive evidence along with oaths.

One of the main objects of Enlightenment criticism of the judiciary was the unsatisfactory treatment of evidential problems. More often than not the judge's discretion amounted to favouritism and corruption. The development of the principle of free evaluation of evidence, combined with systematic burden-of-proof rules, expresses the ideology of Montesquieu and his followers: free, impartial and objective judicial decision-making; but its realization presupposed institutional and scientific development, including a professional police force and forensic medicine.

The rules about the burden of proof developed, at least in the most important Civil Law countries, under the auspices of the *Begriffsjurisprudenz*, starting from the assumption that it was possible to create harmony between the network of systematic relationships among juristic concepts and the goals of legal norms. The development entailed a further step. It was thought possible to create, by means of burden-of-proof rules, harmony between legal systematics and truth.

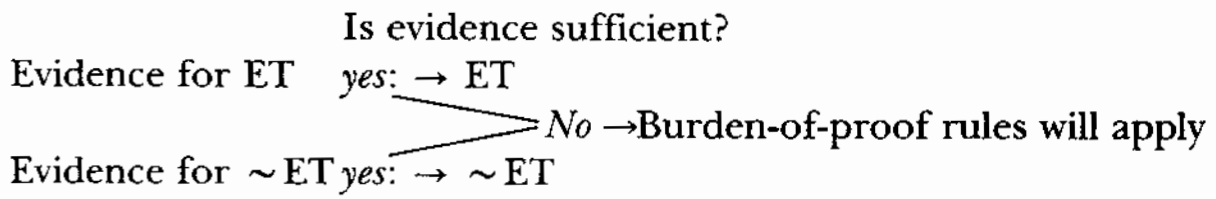
The legal positivism of *Begriffsjurisprudenz* differed, however, from

[3] The text contains our interpretation of the discussion presented, e.g., in the monographs of Leo Rosenberg (see previous footnote) and of Jouko Halila, *Todistustaakan jaosta* (On the Incidence of the Burden of Proof), Porvoo-Helsinki 1955. The interpretation of the ideas of the *Begriffsjurisprudenz* is based on Hannu Tapani Klami, *Länsimaisen oikeusfilosofian historia* (A History of Western Legal Philosophy), 3rd ed. Turku 1980, pp. 175 ff.

the French Exegetical School at one important point. The Exegetical School thought that while statutes constitute the Law, the general principles of Natural Law were the ultimate source should positive sources and methods fail. In *Begriffsjurisprudenz* the conceptual argument was always available, and new law could be created on the grounds of the systematic relationships. Goals of norms did not belong to positive law but rather to legal philosophy (as legal politics was then often called).

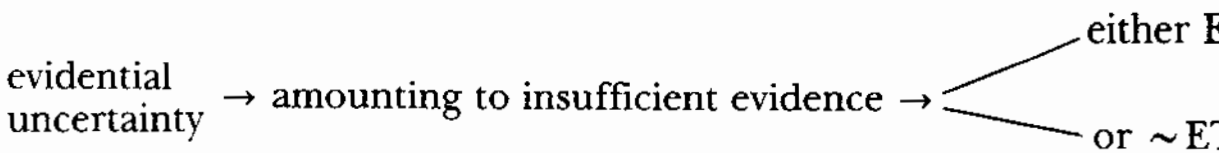
It is only proper to recall some of the background to the development of procedural law, and especially to the burden-of-proof rules. Although the basic ideas of *Begriffsjurisprudenz* are now widely rejected, they have produced a huge mass of rules about the burden of proof. Such rules exist also in Common Law where, however, the main principle in civil cases is a "fair preponderance of evidence" for the evidentiary theme (ET or ~ET, respectively).[4]

The basic idea underlying rules on burden of proof is quite simple: even if the concrete evidence is inconclusive, i.e. insufficient, burden-of-proof rules provide a good, or at least acceptable, solution. The function of the rules can be schematically presented as follows:



Burden-of-proof rules always go in a certain direction, either for or against the ET. If the factual premise to be chosen is the ~ET, i.e. the defendant has the burden of proof, one often speaks of a "reverse" burden of proof. The structure of the burden-of-proof rules is then:

The *scope* of the pertinent rule, defined with the help of legal systematics and juristic concepts — The *choice* of the *factual premise*



[4] On the interpretation of the term "a fair preponderance of evidence", see, e.g., Rupert Cross, *Evidence*, 4th ed. London 1974, pp. 93 ff. There may be different criteria of "fairness" for different situations. The criterion of sufficient evidence is therefore not unitary.

One of the main ideas behind this scheme is the Humean/Kantian ontological difference between Is and Ought, cognitive and normative reasoning. Evaluation of evidence is a matter of cognition; the application of the rules about the burden of proof is a matter of normative reasoning. One should keep them apart, at least if possible. But what about the criterion of sufficient evidence? Does it belong to the cognitive or to the normative part of the reasoning? Here opinions differ, and closer scrutiny shows that many of the main ideas behind the burden-of-proof rules are philosophically *and practically* highly problematical.

According to one view, the criterion of sufficient evidence is a matter of subjective conviction about truth, i.e. a cognitive criterion.[5]

Another view is that evaluation of evidence results in *probabilities* about the ET and ~ET. The evaluation is a purely probabilistic process with, consequently, no place for normative argumentation. Then the decision-maker has to compare the result of his probabilistic reasoning with certain *normatively-based* criteria about the sufficiency of the evidence. If there are no clear norms, he has to set them forth himself.[6]

The presuppposition that burden-of-proof rules should apply is in *both* cases that there is no sufficient evidence for the ET or the ~ET. But the criterion may differ according to either view.

Sufficiency	Criterion	
cognitive—subjective	flexible	fixed
normative—objective	flexible	fixed

It is perfectly conceivable to let the subjective criteria be flexible but, on the other hand, to try to lay down certain fixed criteria that, however, are perhaps interpreted differently by different decision-makers. In Common Law the decision of a juror is in principle a matter of his subjective conviction—but he is told that in criminal cases the guilt of the defendant must be "beyond reasonable doubt".[7]

On the other hand, a normative-objective approach to the sufficiency

[5] This is the main position for example in Finland. See, e.g., Halila, *op.cit.*, pp. 1 ff., and Tauno Tirkkonen, *Suomen rikosprosessioikeus* (The Finnish Law of Criminal Procedure), II, Porvoo 1972, pp. 129 ff.

[6] Thus is the sufficiency of evidency envisaged in Sweden; see Per Olof Ekelöf, *Rättegång* (The Trial), IV, 5th ed. Stockholm 1982, pp. 80 ff.; this view is shared by critics of his evidentiary value method, see, e.g., Per Olof Bolding, *Beviskrav och bevisbörda* (Sufficient Evidence and the Burden of Proof), Lund 1983, pp. 12 ff. Bolding is an adherent of the preponderance-of-evidence principle (*överviktsprincipen*).

[7] What is a "reasonable doubt"? See Cross, *op.cit.*, pp. 97 ff.

of evidence may admit that the criterion (in German *Beweismass*)[8] may vary, depending on different circumstances (for example, one of the parties has caused the uncertainty by destroying relevant evidence; for this reason the normal criterion for similar situations will no longer apply).

But one may go a step further by saying that not only the criterion of sufficient evidence should be elastic but also the burden-of-proof rules. This is because the degree of certainty or probability required for taking the ET or the ~ET as the factual premise of the decision is intimately connected with the application of the burden-of-proof rules. If the criterion is, for example, 80 % for the ET and ⩽20 % for the ~ET, then the plaintiff has the burden of proof for the ET. If one allows an elastic criterion, say, 55, 60, 80, 85 % depending on circumstances, why not admit the possibility of the figure switching over from 60 % to 45 % so that the burden-of-proof rule becomes "reverse", too.

In the Scandinavian countries it is usually assumed that the burden-of-proof rules are *per se* fixed, even though the criterion of sufficient evidence may vary slightly in different situations covered by the same rule.[9] However, a poll inquiry by our project group has shown that practising lawyers were not content with the separation between the evaluation of evidence and the burden of proof.[10] They felt that these influence each other. There was in Finland a clear tendency to consider the concrete evidential situation when establishing the content of the burden-of-proof rules to be applied. In Sweden, practising judges replied that their evaluation of evidence is actually influenced by the burden-of-proof rules. In theory this should not be so, because it entails mixing cognitive and normative reasoning. But we have cause to believe that the theory is wrong here, not the practising lawyers.

In this paper we try to develop a theory of the *functional coherence of evidentiary reasoning*. We expect to meet the objection that the theory does not correspond with valid Swedish or Finnish law. But what is the criterion of "validity" when one strives towards a realistic, rational and

[8] See Dieter Leipold, *Beweismass und Beweislast im Zivilprozess*, Berlin and New York 1985, esp. at pp. 6 ff.

[9] See for Sweden Ekelöf, *op.cit.*, p. 80; for Finland Halila, *op.cit.*, pp. 292 ff.; for Norway Anders Bratholm and Jan Hoel, *Sivil rettergang*, Oslo 1973, pp. 282 ff.; for Denmark W. von Eyben, *Bevis*, Copenhagen 1986, pp. 21 ff.

[10] The results of the survey of our project group "Law and Truth", conducted in Sweden by Mikael Marklund and in Finland by Marja Rahikainen, are presented—apart from these authors' mimeographed cand.jur. theses—in Hannu Tapani Klami *et al.*, "Ett rationellt beviskrav—teori och empiri" (A Rational Criterion for the Sufficiency of Evidence—Theory and Practice), *SvJT* 1988.

functional theoretical model: legal-scientific assertions or the understanding of practising judges?

Before we can actually start building a model, the philosophical and systematic assumptions of the traditional burden-of-proof doctrine and its relationship with legal policy must be scrutinized more closely.

First, assessment of evidentiary values is not exclusively a matter of cognition.[11] It includes—due to the manner in which legally relevant facts are expressed in legal language—many elements of *interpretation and evaluation*. Legally relevant facts are not "brute" facts but to a great extent only understandable in their *normative and institutional context*. For this reason the traditional distinction between questions of fact and questions of law, which goes back to Greek rhetoric, perhaps to Hermagoras from Temnos, is only a relative one. When the evidentiary theme is the danger of certain conduct, negligence, causation, self defence, it is, due to the intertwining of the pertinent questions and the interplay of the evidentiary facts adduced, practically impossible to keep purely cognitive reasoning separate from evaluative reasoning. Hence we recommend defining the notion of evidentiary value in terms of *(un)certainty* rather than in terms of truth and probability. But it must be admitted that legal concepts expressing the relevant evidentiary themes are different in this respect.

Secondly, there is some unclarity about the *conceptual-systematic background* of burden-of-proof rules. Legal rules in general involve a *transformation from teleology to normativity*. In our scheme, teleology includes a cognitive part, namely assumptions about the past, present and future and an evaluation of these. The assumptions include awareness of means-goals relationships, while to the evaluation pertain appraisals of these relationships, *inter alia* from the point of view of expected *cost and benefits*. It is in the nature of this process that knowledge is often uncertain, and for this reason some kind of *strategy* is needed.[12]

Now, there have been different attempts at laying down the conceptual criteria for burden-of-proof rules. Initially, burden-of-proof theories tried to set forth very general criteria, such as facts *constituting or withdrawing* rights, the burden of proof pertaining in principle to the first ones. Once such facts, for example the concluding of a *per se* valid

[11] The problems of the relationships between the structure of language are treated in greater detail in Hannu Tapani Klami, Johanna Sorvettula and Minna Hatakka, "Legal Language and Evidence", to be published in *Festskrift für Kaszimierz Opałek*, Duncker & Humblot 1990.

[12] See Hannu Tapani Klami, "Legal Justification and Control", *Law and Philosophy* 1986, pp. 199 ff.

contract, were proved, however, the burden of proof for invalidating the contract shifted to the party adducing these facts. This was a slightly modified version of the Roman Law principle about the burden of proof of the plaintiff for the presuppositions of getting a formula (*iudicium*; *actio directa*) and the defendant for his defence (*exceptio*; *iudicium contrarium*). Other variants of this theme are the claim/exception "theory" and different "theories" using the rule/exception scheme.

Nowadays very general distinctions have a bad reputation as the conceptual basis for the "incidence" of the burden of proof.[13] The relevance of broad systematic criteria is, however, not wholly denied. So, for instance, there may be a difference concerning the proof of guilt, depending on whether contract or tort is the legal basis of the obligation. But, generally speaking, burden-of-proof rules have become increasingly *differentiated* despite the survival of certain very old maxims as general rules to which there may be *exceptions* such as "nemo suum iactare praesumitur", a presumption for a contract not being gratuitous, or "quisquis praesumitur bonus", good faith is assumed unless proved otherwise.

The differentiation follows the insight that broad burden-of-proof rules do not do justice to the diversity of the evidential and other situations they are intended to cover. Doubts about wide systematic generalizations express a sound scepticism of the basic assumptions of *Begriffsjurisprudenz*. On the other hand, such doubts show that one has begun to analyse rules about sufficient evidence and the burden of proof in a more *teleological* manner: to understand the basic point. The function of such rules is twofold:

(1) To maximize the number of *substantively correct* decisions (approximately decisions based upon *correct factual premises*) even when concrete evidence is insufficient.
(2) To promote certain *substantive goals*, primarily those underlying the substantive norms falling within the scope of a certain burden-of-proof rule. But there may be other goals, too, for example to influence the conduct of the parties in such a manner as to avoid evidentiary uncertainty or to distribute the risk of wrong decisions in a socioeconomically just and equitable manner.

[13] See Halila, *op.cit.*, *passim*, esp. pp. 76 ff.; Per Olof Bolding, *Bevisbördan och den juridiska tekniken* (Burden of Proof and Juristic Techniques), Uppsala 1952, pp. 86 ff.; Bengt Lindell, *Sakfrågor och rättsfrågor* (Matters of Fact and Matters of Law), Uppsala 1987, pp. 284 ff.

Briefly, one is aiming at the truth and the law. But are these goals compatible? Above all, can they be achieved with *fixed rules* on sufficient evidence and on the burden of proof? Or is the most advisable method to *shift the goal and adjust the means* depending on circumstances, i.e. to refrain from inelastic rules and resort to more casuistic reasoning? When shooting, compromises are hardly a good method. A man who wants to kill two birds with one stone will probably hit neither by aiming between them.

Some theoreticians say that the best method for achieving the different relevant goals is in the last analysis to rely on the *substantive truth*, i.e. on the *preponderance of evidence*, even if this is very slight and the evidentiary values for both the ET and the ~ET weak.[14] If any burden-of-proof rules are needed, they should be based upon the *original probability* of the ET and the ~ET in the light of *general experience of life*. In approach the outcome of the evaluation of evidence is a function of truth—admittedly weak—yet the eventual resorting to burden-of-proof rules is a matter of truth, too. One thinks that *wrong decisions cannot promote any justifiable goals whatsoever.*[15] To maximize the number of right decisions is thought the best policy.

On the other hand, it is also maintained that

(1) the risk of *wrong decisions* and the concretization of this risk can be a healthy device in promoting *both* the *general* truthfulness of the factual premises of the decisions *and* the goals of the substantive norms because, admittedly, wrong decisions due to omissions of the parties may have a certain regulatory effect;[16]

(2) the need for fixed rules on sufficient evidence and the burden of proof is so great as to compensate the disutility involved in the risk of a wrong decision.

We think, however, that it is possible to reconcile these seemingly diametrically opposed positions. But this calls for analysis of the *function*

[14] This is in Sweden the position adopted by Bolding, *op.cit.* in footnote 13, pp. 88 ff. and *passim*.

[15] See the brilliant but somewhat one-sided analysis of Lindell, *loc.cit.*

[16] Thus Karl Olivecrona in his much-debated book *Bevisskyldigheten och den materiella rätten* (Burden of Proof and Substantive Law), Uppsala 1930, *passim*; in a similar manner Per Olof Ekelöf, whose most recent writings on this subject are "En rättsvetenskaplig tragedi" (A Scholarly Tragedy) in *Festskrift till Henrik Hessler*, Stockholm 1984, pp. 131 ff., and a review of Bengt Lindell, *Sakfrågor och rättsfrågor*, in *SvJT* 1988, pp. 23 ff. The most radical representative of this view is, however, Henrik Zahle, who also bases his theory of evaluation of evidence on the evaluative appraisal of the conduct of the parties; see his *Om det juridiske bevis* (On Juristic Evidence), Copenhagen 1976.

of the rules about sufficient evidence and the burden of proof. The next question will therefore be the following:

3. IS THERE A NEED FOR FIXED RULES AT ALL?

The question is difficult because there is practically no empirical evidence to back the assumptions about the function of such rules. The situation is paradoxical, to say the least. There is discussion of the function of burden-of-proof rules and there is also discussion of the notion of sufficient evidence. But it is not clear *who* has to *prove* his assumptions, or to what extent!

It has been said that *fixed* rules about the burden of proof are backed by certain arguments of legal policy of their own:

(1) They promote the achieving of the substantive truth, because they are at least partly based upon general experience of life, and

(2) because they compel litigants to secure evidence beforehand and to present it in court.

(3) Unnecessary litigation is avoided, because the parties and their counsellors can evaluate the evidence beforehand and thus rightly estimate their chances in a lawsuit.

(4) It is, apart from (3), important to be able to foresee the verdict of the court in an actual lawsuit.

(5) Clear burden-of-proof rules will help the judges (or the jurors) when laying down the factual premises for their decision.[17]

But do these arguments hold good, and how do they relate to *uncertainty* regarding the factual premise(s) of the concrete decision? We are again faced with the problem of whether rules about uncertainty in decision-making really affect the pre-trial conduct of the parties, in particular the securing of evidence.

Our empirical survey indicates that Swedish and Finnish lawyers operate with very high criteria for the sufficiency of evidence, both in civil and in criminal cases.[18] Clearly, this will often lead to application of burden-of-proof rules. But the lawyers also realize the reverse: the higher the required certainty about the evidentiary theme is, the higher is also the risk of wrong decisions. But when the lawyers were asked how frequent they thought wrong decisions were, their answers varied considerably. No wonder, because one never does know how often a court

[17] See the writers quoted in footnote 15 and Halila, *op.cit.*, pp. 76 ff.

[18] See Klami *et al.*, *SvJT* 1988.

makes a mistaken evaluation of evidence or a wrong decision due to uncertainty and the ensuing application of the burden-of-proof rules. That the number of faulty convictions was estimated to be low (although the general ratio of total acquittals in criminal cases is $<5\%$) was not a surprise; nor was it surprising that the ratio of wrongful acquittals was thought to be high, because these are interdependent. (There was, however, in the different individuals' ratings no such functional coherence as the theory of probability would suggest.) More problematical were the guesses that wrong decisions in either direction were rather common in civil cases, especially where a justified claim fails because of lack of evidence.

Now, given that the design of the burden-of-proof rules has become to a great extent teleological, aiming at influencing the evidentiary behaviour of the parties and securing the substantive goals of the norms, is it just and equitable even to *try* and promote general goals with an uncertain number of wrong decisions in concrete cases? May one really say to a plaintiff who has lost his case because of rigid burden-of-proof rules: "Well, your claim was *probably* justified, but we had to sacrifice the truth—and your claim—to achieve certain general goals of law?"[19] This sounds odd indeed, and the point of the criticism of rigid burden-of-proof rules is exactly this: goal-directed rules may entail wrong, unacceptable decisions.

But discussion of the risks and disutilities of wrong decisions remains speculative because there has been (a) no information about their *frequency* or the possible *impact* of different burden-of-proof rules on it, and (b) there is no *objective* measure of the disutilities involved in wrong decisions.

To discuss the need for fixed burden-of-proof rules we will try to develop a model for assessing the disutilities and the ensuing requirements for certainty. But this must be preceded by analysis of the teleology behind the rules.

4. TRUTH OR GOALS?

Four main arguments have been used for or against different burden-of-proof rules: we exclude purely *conceptual arguments* whose relevance

[19] A brilliant criticism of this type of reasoning is included in Bolding, *Beviskrav och bevisbörda*, pp. 24 ff. See also Lindell, *op.cit.*, pp. 284 ff.

depends on basically *substantive reasons* since their function is to relate the use of substantive reasons to consistent criteria.[20]

The main arguments can be grouped as they relate to

(4.1) *the truth*:

(4.1.1) *the original probability* (OP) of the relevant facts in the light of general experience of life

(4.1.2) the reasons for and the possibilities of securing evidence in advance and of presenting it in a lawsuit: the *evidential possibilities* of the parties (EP)

(4.2) *the goals*:

(4.2.1) the *different goals behind the substantive norms*—the norm goals (NG). The goals may converge so that the concrete facts/legal consequences relationship expresses a means-goals assumption pertaining to the "vector product" of the goals rather than to a single goal; also, there may be a certain scope for a choice between different legal consequences. Nevertheless, rules about the burden of proof and about sufficient evidence should normally function together with the substantive norms and secure their effectiveness.

(4.2.2) When one admits that wrong decisions are possible, one will also ask which of the parties typically or in a concrete case can better bear the socioeconomic consequences of a faulty decision; such "rightness reasoning" about the just and equitable "incidence" of the risk of wrong decision can be called the *risk aspect* (RA) of sufficient evidence and the burden of proof.

Our idea is that evidential decision-making follows the normal principles of decision-making theory. The problem is *uncertainty*, but not only evidential uncertainty. There may be uncertainty about the goals themselves—the decision-maker (or the legislator or a legal scholar) feels uncertainty for instance about what the goals of a certain norm really are.

Moreover, there may be uncertainty about the mutual relationships

[20] The analysis of the arguments is based on legal writings on the burden of proof: we have used mainly as references the books of Rosenberg, Halila, Bolding and von Eyben. Since we are not discussing concrete problems, we refrain from giving concrete references. The reader is probably aware that we refer mainly to Scandinavian legal writings, not because we think that they represent the avant-garde of legal science in the law of evidence (although in the field of evidentiary value Professor Ekelöf has done much pioneer work). The reason is simply that our model is based partly on an empirical survey of Swedish and Finnish lawyers who have read mainly domestic treatises and monographs. To understand the results, knowledge of their doctrinal background is necessary.

between the truth problems and the goals. For example: how will uncertainty about facts affect achievement of the norm goals (NG)? Or, because one does not know for sure the reasons for or the extent of evidential uncertainty, it may be difficult to say anything definite about the just and equitable incidence of the risk. In short, the relevant arguments may in theory:

(a) *concur*, e.g.: pro: OP, EP, NG, contra: unclear RA
(b) *contradict* each other, e.g.: pro: OP, contra: RA
(c) be *intertwined* in such a manner that uncertainty about one argument may make another argument "fuzzy" (to use a term common in decision-making theory).

In such a situation it is wholly understandable that there are differing opinions concerning the rules on sufficient evidence and the burden of proof. For example: original probability goes against the weaker party, but the stronger party would no doubt be more able to bear the consequences of a possible decision. Moreover, one may say that the stronger party, generally speaking, has better possibilities to secure evidence in advance or to present it in court. Hence, uncertainty should also be an argument for the burden of proof being carried by the stronger party.

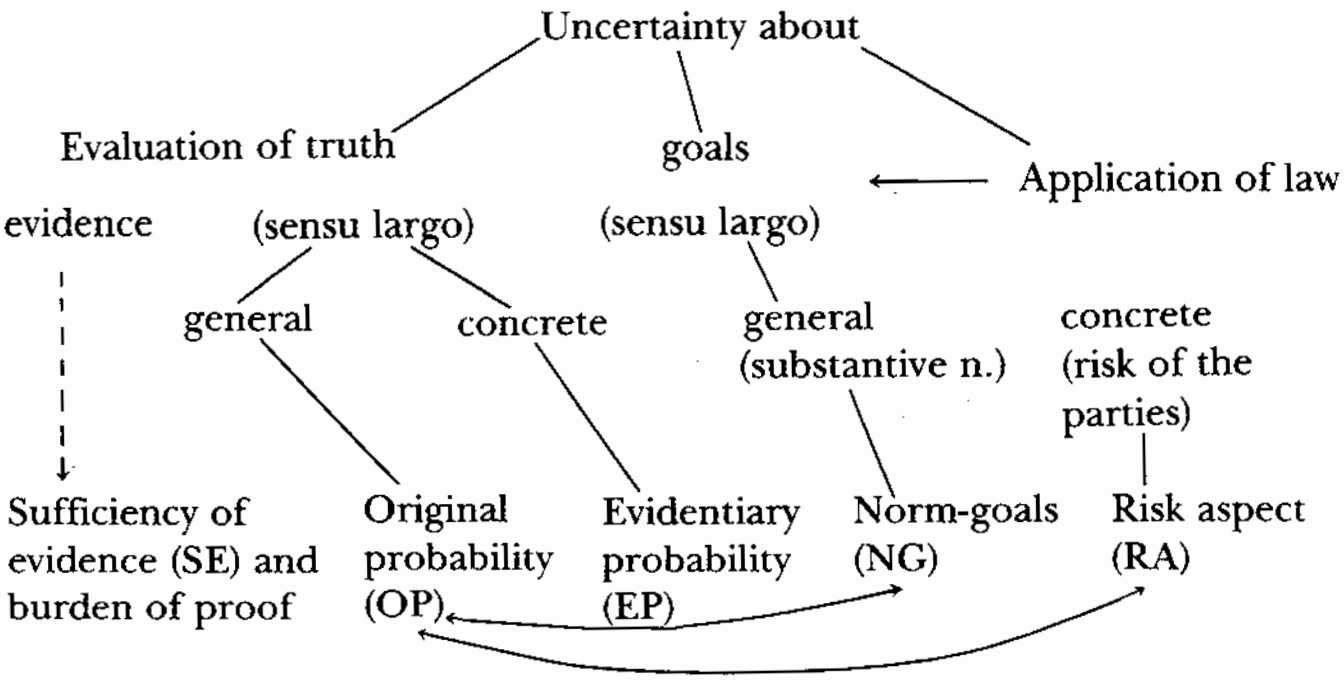
In criminal cases the *risk aspect* may be so heavily stressed that, e.g. in cases of rape, in spite of the high original (general) probability of evidence (the women's reports) and the securing of the goals of the norms (the protection of the sexual integrity of women) it may call for great caution when convictions are concerned. Securing evidence in advance or presenting it in court may be equally problematical for both parties.

Although most debts are duly paid and although debtors are usually weaker socioeconomically than creditors, the EP aspect strongly indicates that the debtor should bear the burden of proof of payment. Hence most legal orders have assumed that such a rule goes against the debtor. It is, after all, fairly easy to secure evidence with a receipt; but it remains somewhat problematical, whether compelling a concrete poor debtor to pay twice will really affect the behaviour of *debtors in general*, at least to a significant extent.

One perplexing problem is that the definitions of the arguments may be somewhat different, depending on the attention focused here upon the *criterion of litigation*. For example: rapes are so unusual that the OP of a man selected at random of being guilty of rape is very low; on the other hand, one must assume that the reports of a woman against a

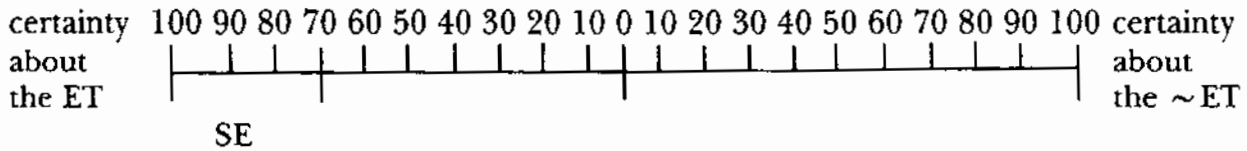
certain man are highly reliable, generally speaking. Most debts are paid in a due manner, but it is probably rather uncommon that a creditor duns a debtor who has already paid, in order to get fraudulently paid twice. This, again, is because debtors have usually taken a receipt, and this requirement may, again, depend on their knowledge of the burden of proof.

But despite the difficulties it appears possible to develop a model for rational reasoning about sufficient evidence and the burden of proof. Our starting-point is the following scheme:



(The connecting arrows are only examples)

We consider that the criterion of sufficient evidence is a function of these four arguments, i.e. original probability, evidentiary possibility, the norm goal and the risk aspect. If one is able to establish the point of sufficiency on the probability/uncertainty scale, this automatically means agreement upon the content of the burden-of-proof rule.



If the SE for the ET = 70 %, then (a) the SE for the ~ET ⩽ 30 % and, consequently, (b) there is a burden of proof about ET, because in these contexts one has to start from the two-way axiom of probability $P(T) \leftrightarrow 1 - P(T)$.

5. UNCERTAINTY, THE DISUTILITY OF A WRONG DECISION AND THE CRITERION OF SUFFICIENT EVIDENCE

The evaluation of evidence often results in some uncertainty about the ET and the ~ET. Complete certainty about the past or the future is often difficult to achieve.

Here one should apply the *evidentiary value method*[21] which starts from the one-way notion of probability, i.e. $P(ET) + P'(\sim ET) \leqslant 1$, because this focuses attention upon the probability (in our opinion certainty is better) of the evidence proving the ET and not, as the theme method does, upon the probability of the theme of the information, i.e. P(ET/EF), where EF means Evidentiary Fact, such as the observations of a witness.

The evidentiary value method asks whether certain criteria of relevance for the relationship between the evidentiary facts and the ET are met. The method presupposes the existence of an experience-based *evidentiary mechanism*[22] that has worked in a concrete case. The standard

[21] We have dealt with the evidentiary value method in several papers containing references to the relevant legal writings: Hannu Tapani Klami, *On Truth and Evidence*, Turku 1986, esp. at pp. 42 ff.; Hannu Tapani Klami, Marja Rahikainen and Johanna Sorvettula, *Todistusharkinta ja todistustaakka* (Evaluation of Evidence and the Burden of Proof), Helsinki 1987, pp. 24 ff. and *passim*; Hannu Tapani Klami, Johanna Sorvettula and Minna Hatakka, "Original Probability and Evidentiary Value", to be published in A. Aarnio and C. Varga (eds.), *A Finnish-Hungarian Symposion on the Rule of Law*, 1990; *id.*, *Evidentiary Value* (in publication). But we refer here to some of the writers quoted by us: Per Olof Ekelöf, *Rättegång* IV, pp. 7 ff.; *id.*, "Beweiswert" in *Festschrift für Fritz Baur*, Tübingen 1981, pp. 343 ff.; *id.*, "My Thoughts on Evidentiary Value", in P. Gärdenfors *et al.* (eds.), *Evidentiary Value, Essays Dedicated to Sören Halldén*, Lund 1984, p. 9; Sören Halldén, "Indiciemekanismer" (On the Mechanisms of Circumstantial Evidence), *TfR* 1974, pp. 55 ff.; Anders Stening, *Bevisvärde* (Evidentiary Value), Uppsala 1975 (with an extensive summary in English); Robert Goldsmith, "The Evaluation of Evidence by Swedish Judges and Public Prosecutors", in *Acta Psychologica* 1982; see also the essays by Peter Gärdenfors, Bengt Hansson and Nils Eric Sahlin in the volume *Evidentiary Value*. In the United States there have been a few adherents of the method, e.g. Glenn Shafer, *A Mathematical Theory of Evidence*, Princeton 1970, and in West Germany recently Helmut Rüssman, "Zur Mathematik des Zeugenbeweises", in *Festschrift für Heinrich Nagel*, Münster 1987, pp. 332 ff. The standard work on judicial probability is L. Cohen, *The Provable and the Probable*, Oxford 1977. For a criticism of the evidentiary value method, see, e.g., Zahle, *Om det juridiske bevis*, pp. 280 ff.; Per Henrik Lindblom, review of Stening, *Bevisvärde*, in *SvJT* 1977, pp. 280 ff.; Bolding, *Beviskrav och bevisbörda*, pp. 12 ff.; Lindell, *Sakfrågor och rättsfrågor*, pp. 144 ff. (but see the reviews of this book by Ekelöf and Nils Eric Sahlin in *SvJT* 1988, pp. 21 ff., and by Hannu Tapani Klami, *FJFT* 1989, pp. 72 ff. ; Torstein Eckhoff, "Temametode eller verdimetode i bevisvurdering" (The Evidentiary Value Method or the Theme Method in the Evaluation of Evidence), *SvJT* 1988, pp. 321 ff.

[22] On the concept of an evidentiary mechanism, see Halldén, *TfR* 1974, pp. 55 ff.; Martin Edman, "Adding Independent Pieces of Evidence", in *Essays Dedicated to Sören Halldén on his 60th Anniversary*, Lund 1974; Stening, *Bevisvärde*, pp. 34 ff. But see our criticism in Klami-Sorvettula-Hatakka, "Evidentiary Value", *FJFT* 1989, pp. 23 ff.

case is that the ET has been caused by the theme, such as in the case of the witnesses' testimony (EF) being based on, "caused by", observations about the crime taking place (ET).

Now, if the evidentiary mechanism has not worked, the EF tells us nothing about the ET which, nevertheless, may be true or false. The evidentiary value is defined as the probability of the evidentiary mechanism having worked. If it has not worked, we know (on grounds of the EF) nothing exact about the ET. For this reason we cannot conclude $P(ET/EF) \rightarrow 1 - P(\sim ET/EF)$, where (ET/EF) means evidentiary theme given evidentiary fact.

The evidentiary value method does not deny, however, the possibility of drawing conclusions from the probability of the ET to the (im)probability of the $\sim$ET; but this possibility must be based upon the conclusive nature of the information in question and not be taken axiomatically for granted.

We will not here enter into the details of the evidentiary value method. Its peculiar feature is the unilateral certainty/uncertainty conception which allows for the evidentiary values for both the ET and the $\sim$ET being rather low. Although we think that the evidentiary value method has certain weaknesses due to its somewhat ultra-positivistic background assumptions (we have in another paper suggested certain revisions), at this point it is realistic enough. After all, evidence may be weak in both directions, and it would be incorrect to assume that weak evidence for the guilt of the defendant should automatically constitute strong evidence of innocence.[23]

But the court may nowadays no longer say "non liquet" and refrain from making a decision, like a Roman private *judex* could do; neither can it postpone its decision indefinitely, until new evidence should be found, as the old Swedish-Finnish criminal courts could do. For this reason rules on sufficient evidence and the burden of proof start from a *bilateral notion* of certainty or uncertainty. If the evidence for the ET is insufficient, the factual premise of the decision will be $\sim$ET—but neither is this the same as saying that there should be sufficient evidence for $\sim$ET. The choice of $\sim$ET as the factual premise is not based upon its evidentiary value but upon *law*. For this reason e.g. *Tybjerg* has

[23] This inference was almost unanimously rejected by the subjects of the poll survey made in Sweden. In Finland the question was somewhat different but the results were similar: one preferred the evidential value of the evidence to the unspecified probability of the theme.

characterized the resorting to rules about the burden of proof as an announcement of the bankruptcy of the judiciary.[24]

But is it necessary to make a choice between law and truth?

We suspect that the basic position of the burden-of-proof theory sketched above needs a certain revision, but this need not result in a complete rejection of rules and recourse to reasoning about (the fair) concrete preponderance of evidence. We now present the model and shall later comment on the possibilities of using it.

The starting-point is the formula of decision-making theory regarding the certainty (probability) needed for making a certain decision under uncertainty. When a *linear strategy* is adopted, the minimum probability is given by the following formula, where $P_{EV/min}$ denotes the minimum probability (here: evidentiary value) for convicting the defendant, when D_g = the disutility of a wrongful acquittal and D_i = the disutility of convicting an innocent defendant:

$$P_{EV/min} > \frac{1}{1+\frac{D_g}{D_i}}$$

If $D_g = 20$ and $D_i = 100$ at a scale from 0 to 100, our formula gives the following figure:

$$P_{EV/min} > \frac{1}{1+\frac{20}{100}} = \frac{1}{1,2} = 0.0833 = 83.3\,\%$$

This formula—in evidentiary contexts first used by Kaplan[25]—is based upon certain assumptions. First, the starting-point is the two-way probability according to which $P_{EV/min/ET} = 1 - P_{EV/min} \sim ET$. But, as we will see, this does not exclude the assumption about *inconclusive EVs*, $P_{EV/ET} + P_{EV/} \sim ET \leqslant 1$.

Secondly, it is assumed that *correct decisions involve no disutility* whatsoever. To this may be objected that correct decisions actually may cause

[24] E. Tybjerg, *Om Bevisbyrden* (On the Burden of Proof), Copenhagen 1904, p. 1.

[25] See J. Kaplan, "Decision-Making Theory and the Factfinding Process", *Stanford Law Review* 1968, pp. 1072 ff. Our application and further elaboration of the formula was first presented in Hannu Tapani Klami, Marja Rahikainen and Johanna Sorvettula, "On the Rationality of Evidentiary Reasoning, a Model", *Rechtstheorie* 1988/3. The present essay contains certain revisions of our original positions, especially as far as the burden-of-proof problem is concerned.

at least certain inconvenience. It is well known that *per se* justified punishments may involve disutilities for the convicted person, for his family and for society, disutilities that may well surpass the (alleged) positive effects of the punishment. And this is so even if the conviction should be based on correctly assessed facts and due application of legal norms. But this issue has nothing to do with *evidential* uncertainty; it is irrelevant from this viewpoint and can therefore be left out.

The third assumption is that there is *one unitary utility*—or in this case *disutility*—over different decision alternatives. *This* assumption needs certain revision. We have seen that there are arguments pertaining to *truth* and to *goals*, OP, EP, NG and RA, and that they may concur, contradict or be intertwined. In such a situation the assumption about one unitary disutility is *too undifferentiated*. Assessment of the disutilities of wrong decisions in either direction should somehow express this concurrence or contradiction. Otherwise it would be difficult to discuss the figures in an explicit manner: too many aspects would remain hidden behind the curtain of intuitive, holistic reasoning. And it is important to realize that in spite of the intertwining of different arguments there is in legal reasoning no unitary utility. The presence of *goal arguments* in discussion about the burden of proof may be considered as more justified or less. But wholly unjustified it is certainly not, and this indicates that evidentiary reasoning is in the last analysis a special case of decision-making under uncertainty, presupposing a certain *strategy* within the terms of assessing *disutilities*. Truth is a matter of cognition. But when one has to make decisions and there is uncertainty about the truth, the picture will change somewhat. One has to evaluate the uncertainty and the ensuing risk of wrong decisions.

We suggest that the four relevant arguments used when discussing rules about sufficient evidence and the burden of proof should in concrete decision-making situations be made explicit and used as the criteria for assessing the disutilities of possibly wrong decisions. The arguments are: OP, EP, NG and RA.

This method makes the relevance of the different arguments explicit: if they concur, the disutility values for the D based on the factual premise(s) ET and ~D/~ET converge. Should they contradict, the disutility values will diverge when assessed with the help of the arguments.

There is admittedly one defect: a *separate* evaluation of the disutilities of wrong decisions fails to pay attention to the possibility that the relevant arguments *intertwine*. But we think that our next step will at least partly solve the problem: the arguments should be hierarchically

ordered and weighed against each other, the assumption being that this stage of reasoning will express at least partly the relevant interdependence of the different arguments. But there are certain problems as to the use of the OP, EP, NG and RA as criteria of the possible disutility of wrong decisions.

First, some words about the *disutility scale*. It may very well be 0–100, because one need not assess the general disutilities of all kinds of wrong decisions, only the disutilities involved in wrong decisions about the *concrete* decision alternatives, related to each other from this point of view. In this context the measuring of disutilities, related to each other, is by no means hampered by the fact that the ETs of the parties are not necessarily diametrically opposite. The only presupposition is the ability to relate the different decision alternatives in a concrete case *to each other* with the help of the scale and that the scale used is, regarding *disutility*, somewhat unitary, i.e. that a disutility of 50 units from the point of view of OP corresponds to an equal disutility, measured in terms of RA. It is in theory unimportant how the figures of the scale are chosen, but we recommend 0–100 for the sake of simplicity. Because only the relationship D_g/D_i counts, it does not matter what figure distribution is used.

Should there be more than two decision alternatives, for example conviction for murder, conviction for manslaughter, acquittal, one may assess their disutilities separately and compare them to each other pairwise: D1/D2, D2/D3.[26]

The method does not presuppose that the question should pertain only to conviction or acquittal. Acquittal may be wholly excluded, the choice concerning e.g. the type of crime that has been proven/not proven, or different decision alternatives in a civil lawsuit.

When the disutilities are assessed, one need not presuppose that $D_g + D_i = 100$. Both disutilities may be estimated low—but it is their relationship that counts, e.g.: $20/30 = 40/60 = 60/90$.

We are, however, aware of the theoretical and practical difficulties involved in measuring disutilities in the light of arguments OP, EP, NG

[26] It is interesting to note an American experiment where the subjects were asked to decide about the guilt of the defendant on grounds of documents from an actual trial. Their reluctance to convict depended significantly on the choices given to them: in a situation first-degree murder/acquittal, the percentage of acquittal was much higher (54 %) than when the choice pertained to a second-degree murder/acquittal (8 %). On this experiment of Neil Vidmar (whose original paper was not accessible to us), see B. Grofman, "Mathematical Models for Judge and Juror Decision-Making", in B.D. Sales (ed.), *The Trial Process,* New York 1981, pp. 312 ff.

and RA. These are, after all, *not* by their nature grading concepts. Moreover, there may be uncertainty about their content. But it is very important again to note that the method is not an objective method for laying down the criteria for sufficient evidence. It is a method for testing the subjective rationality of decision-making and for facilitating inter-subjective discussion about the conditions of rationality. For this reason one must admit a certain imprecision, different interpretations of the arguments and differences as to their concrete application. But the $P_{EV/min}$-values are to be compared to concrete P_{EV}-values (for ET and $\sim$ET). These values, too, express results of subjective reasoning about certainty. The disutility method aims at establishing the sufficient (subjective) conditions of this certainty.

Before considering concrete problems of assessment of the disutilities we would stress one basic feature of the method. It presupposes that at least *some* disutility is involved in making a wrong decision in either direction. Thus where D_g is estimated $=0$, the formula will presuppose *full certainty*, irrespective of the value of D_i:

$$\frac{1}{1+\frac{D_g}{D_i}} = \frac{1}{1+\frac{0}{x}} = \frac{1}{1} = 100\,\%$$

On the other hand, if $D_i=0$, the formula does not work at all because, irrespective of the value of D_g, >0, $x/0$ being equal to ∞,

$$\frac{1}{1+\frac{D_g}{D_i}} = \frac{1}{1+\frac{x}{0}} = \frac{1}{1+\infty} = \frac{1}{\infty}\,!$$

Now, when in our empirical survey certain subjects answered that conviction in criminal cases presupposes 100 % certainty (or even more than so!), they apparently did not fully realize the logic of this requirement: with the disutility formula full certainty presupposes that no disutility whatsoever is involved in a wrongful acquittal. Surely no judge would agree. On the other hand, it would be unrealistic to assume no disutility for wrongful convictions either. More generally, the structure of the formula is realistic insofar as at least *some* disutility is involved in making a wrong decision in either direction, and, consequently, because neither D_g nor D_i can in civil or criminal cases be $= 0$, the criterion of sufficient evidence is >0 and <100. This means simply that one cannot be content with no evidence at all, nor can one require full certainty, either.

5.1. *Original Probability as a Measure of Disutility*

The basic idea is that as the disutility involved in a wrong decision increases, so does the original probability against it. This holds at the general level: truthfulness of decisions is also a kind of *goal*, and it is the more endangered, the more decisions under uncertainty are made that go against general experience of life.

But it holds at the concrete level too that the greater the disutility, the more concrete wrong decisions are made against the OP. One may refer to the expectations of the party who is really right and to the fact that a dubious decision made against the OP is felt to be so unjustified as even to endanger the general legitimacy of the judiciary.

5.2. *Evidentiary Possibilities and Disutility*

Here the idea is that the extent of the disutility of a wrong decision depends on whether it would have been possible to prevent the wrong decision by securing evidence and/or presenting it in the lawsuit. We see that the assessment of disutility with the help of the EP has both normative and "cognitive" features. One may—or may not—speak about the *duty* of the parties (at least in a civil case) to secure evidence in advance and to present it in court.[27] But such an abstract "duty" depends on the evidential possibilities of the parties, and this is so even for the weight of this normative argument.

So one has here to evaluate the *evidential conduct of the parties*. The greater the disutility of a wrong decision is, the easier it would have been (or should have been) for either party to prevent it by presenting evidence. Here of course it is possible that neither party can be blamed, or that both have been equally negligent; but if the disutility from this point of view is equal, this part of the evaluation indicates a neutral position:

[27] The burden of proof is often defined in German and Scandinavian writing as a *duty to prove*, an obligation of the parties whose *sanction* is that if something relevant has not been proven, the party in question has to bear the *negative consequences* of the failed proof; see, e.g., Halila, *Todistustaakan jaosta*, pp. 2 ff. The criterion of the sufficiency of evidence is in Swedish called *beviskrav*, literally: evidentiary requirement. But such a definition fails to pay attention to the fact that the law may oblige a party to present evidence also for circumstances that are adverse for his case. This is in Scandinavian law the position of a public prosecutor; as for taxation cases, the taxpayer is under obligation to present truthful information about relevant facts even if they should be detrimental to him.

$$\frac{1}{1+\frac{D_g}{D_i}} = \frac{1}{1+\frac{10}{10}} \left(\text{or, e.g.} \frac{1}{1+\frac{50}{50}} \right) = \frac{1}{1+1} = \frac{1}{2} = 50\,\%$$

i.e. a preponderance of evidence will suffice in such a situation.

It is no bar to this assessment of disutility that the defendant in a criminal case is under no real duty to present evidence about, or to explain, his conduct. Certain evidentiary shortcomings may nevertheless be used against him, for example if he presents a false alibi or destroys evidence.[28] But this position of present law may be directly taken into account when evaluating the disutility of a wrong decision from this point of view. This indicates simply an evaluative trend to keep the disutility value of a possible wrongful acquittal low.

5.3. *The Goals of Substantive Norms and Disutility*

It is theoretically relatively easy to use the NG as a measure of disutility. For most wrong decisions in either direction involve disutility—but the disutilities need not be equal. Certain problems may arise when the decision presupposes application of a complex of norms or when there are several goals behind the norms and wrong decisions in either direction would relate to different goals. But we assume here that one should focus upon the concrete norm or, respectively, upon those goals whose function would be *most* marred by a wrong decision in either direction. Then one should assess the disutilities by weighing the *importance* of the respective norms and/or goals behind them. Lawyers are accustomed to such comparisons.

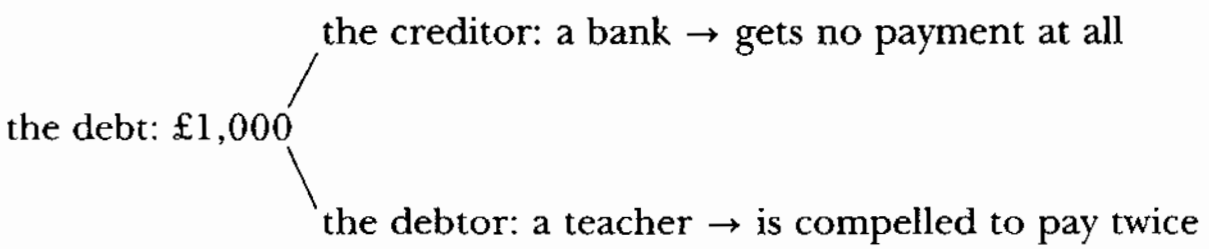
This measuring of disutility also takes into account the concrete nature of the evidential problem. The evidential problem may concern the different parts of the norm, and may thus be somewhat differently evaluated from the point of view of disutility. For instance, in a tort case the evidential problem may deal with negligence, causation or the amount of damage. It is not clear that a wrong decision should affect the goals of the norm(s) in a similar manner in all these situations. (There may, *inter alia*, be differences as to the estimated frequency of evidential problems pertaining to those criteria of responsibility.)

[28] Certain Finnish writers even say that if the defendant in a criminal case refuses to answer a question, this may—in a free evaluation of evidence—be used as circumstantial evidence against him (see Tirkkonen, *Suomen rikosprosessioikeus* II, p. 120). We do not share this view.

One should also refrain from attaching any positive utility to wrong decisions from the point of view of the NG, even though the method allows such an attitude: this can be taken into account when establishing the disutilities.

5.4. *The Socioeconomically Just Risk of Wrong Decisions and Disutility*

Here the assessment of disutility is based on the socioeconomic status of the parties and their ability to bear the consequences of a wrong decision. This is an issue of social justice and equity that does not depend only on the social strength of the parties but also on the concrete nature of the consequences. One may even ignore the social background of the actual parties; this will probably be the case in criminal law. Wrongful conviction for, e.g., murder would involve nearly maximal disutility for the defendant, irrespective of his social background or wealth. Generally, however, this criterion calls for comparisons that relate the consequences of wrong decisions to the socioeconomic status of the parties, for example in a debt case:



The RA criterion should be kept at a rather concrete level, thus expressing the peculiar elements of the evidential/judicial situation. For here one should consider the amount of the debt and possible differences in the parties' social status. The debtor may be a firm that is as sound as the bank; the creditor and the debtor may be almost equally poor, and so on.

Disutilities may be assessed in two ways:

(I) concretely, taking into account the *actual features* of the case
(II) abstractly, focusing upon *typical* features of evidential situations.

But here an "inductive" method is recommended. Discussion of typical disutilities should lead to analytical, concrete assessment of disutilities, in experiments or in judicial practice, and this should precede conclusions as to general disutilities as criteria for rules about sufficient evidence and the burden of proof.

We have already referred to the discussion of the relative strengths of

the different arguments in the construction of burden-of-proof rules. Some lawyers wish to operate mainly with *original probability*, asserting that the general truthfulness of the ET/~ET is the best means for achieving the NG and that it is also the best criterion for the incidence of the risk of wrong decisions made under uncertainty. Certain Scandinavian writers have viewed the rules about sufficient evidence (the burden of proof) mainly as a legal reaction to the "evidential behaviour" of the parties: Karl Olivecrona early expressed such thoughts.[29] Even a heavy burden of proof was, according to him, justified in order to sanction the duty of the parties to secure evidence and present it during the lawsuit. Professor Ekelöf, the founding father of the evidentiary value method, has accepted this course of reasoning, simultaneously stressing that the functioning of substantive norms depends on evidentiary reasoning. Professor Zahle has tried to construct a general theory of evidence, based upon the evaluation of the "evidential conduct" (*adfærd*) of the parties, at the same time, surprising enough, pleading for the theme method with its two-way probabilities.[30] And, more recently, socially oriented theories about the burden of proof are rather common, too, although our own empirical survey indicated that judges were rather reluctant to sacrifice substantive truth in order to protect the weaker party. Some answers explicitly stated that judicial truth is not a matter of social policy.[31]

In any case the method has to allow for different weightings of the measuring arguments, and, consequently, for different weightings of the disutility values assessed. But there is a method for ordering the arguments hierarchically: Saaty's Analytic Hierarchy Process (AHP).[32] This is based on a *pairwise comparison* from the point of view of *importance*. In our case, where there are four relevant arguments, the scheme of the AHP comparison would be as shown opposite.

This method has been used with success in poll studies. It has certain practical advantages—admittedly at the cost of imprecision and subjective inconsistency. Psychological studies of decision-making have re-

[29] See Olivecrona, *Bevisskyldigheten och den materiella rätten*, pp. 153 ff.

[30] Zahle, *Om det juridiske bevis*. pp. 34 ff. and *passim*. But see the acute criticism delivered by Per Olof Bolding, "Bevisvärdering utan sannolikhetsuppskattning" (Evaluation of Evidence Without Assessing the Probability), *TfR* 1978, pp. 530 ff.

[31] See Klami *et al.*, "Ett rationellt beviskrav", *SvJT* 1988.

[32] On this method, see T. Saaty, *The Analytic Hierarchy Process*, New York 1980. The method has been successfully used in Finland by Liisa Uusitalo, *Suomalaiset ja ympäristö* (The Finns and the Environment), Helsinki 1986, in a poll study on the attitudes of Finns towards the goals of environmental policy.

	Is one argument as important as another?	if more important, how many times more?							
OP/EP	(1)	2	3	4	5	6	7	8	9
OP/NG	1	2	3	4	5	6	7	8	9
OP/RA	1	2	3	4	5	6	7	8	9
EP/NG	1	2	3	4	5	6	7	8	9
EP/RA	1	2	3	4	5	6	7	8	9
NG/RA	1	2	3	4	5	6	7	8	9

vealed a tendency to compare goals or arguments exactly in this manner, i.e. pairwise. Moreover, the pair comparison does not presuppose that the different comparisons are wholly transitive, i.e. consistent. For it is possible that a preference between different/goals or arguments is *differently conditioned* when one has to compare pairwise, for example OP/EP, OP/NG and EP/NG. An overall transitivity of the comparisons need therefore not be presupposed.

Moreover, uncertainty about the arguments used for evaluating the disutilities can be made explicit and operational with the AHP method. There may be uncertainty about the *content* of an argument or its *applicability* to the assessment of uncertainty. But in such cases one may use *relatively low weightings* when comparing the "fuzzy" argument to an argument more clearly defined and relevant.

The mathematical analysis of the AHP weightings is simple. If two arguments are considered equally important (= 1), both get the value 1. If OP is considered to be twice as important as EP, the values for OP and EP are 2 and 1/2, respectively. In our scheme, where there are four arguments and three comparisons for each, the total hierarchic value of an argument (OP) is got from the formula

$$\sqrt[4]{\text{OP/EP}\times\text{OP/NG}\times\text{OP/RA},}$$ i.e.
from the $\sqrt[4]{}$ of the product of the different weightings.

It is then easy to transform the $\sqrt[4]{}$ product values into percentages and thus to obtain the *weightings of the disutilities assessed with the help of the OP, EP, NG and RA*. Simple calculations (with a personal computer) will then result in a weighted assessment of disutility that expresses the possible concurrence or contradiction of the arguments used for assessing the disutilities and their interrelated relevance. This method also

pays attention to the possible vagueness of the measure for disutility. In short, despite certain shortcomings the disutility formula

$$P_{EV/min} > \frac{1}{1+\frac{D_g}{D_i}}$$

together with the AHP technique will provide solid ground for assessing the subjectively rational criteria for the sufficiency of evidence and the burden of proof. At the very least this method calls for making all stages of reasoning explicit.

Weighing the arguments may, admittedly, be a difficult task. But in the era of computers there may be certain pre-programmed AHP weightings that are not necessarily binding for the decision-maker but only guide his reasoning when he is assessing the disutilities. For we think that in the future, once it has been realized that evidentiary reasoning involves decision-making under uncertainty, there will appear a need for making the reasoning as explicit and rational as possible. One cannot be content with wholly intuitive decision-making. The lowering of the criteria of sufficient evidence as a means for legal policy is already being discussed. In the sphere of consumer protection and product liability it is not infrequently said that the legislator should intervene in evidential reasoning to secure the protecting function of the substantive norms. However, such legislatory intervention should be preceded by an analysis of the arguments; on the other hand, this presupposes that the courts understand the language and techniques of such a regulation.

If the disutilities are not assessed clearly enough against a certain criterion, the assesser may neutralize the criterion by assessing equal disutility, say, 0/0, 10/10 and so on, for different, possibly wrong, decision alternatives.

The method obviously calls for empirical tests where the weightings of the arguments are *pre-programmed* allowing, however, concrete deviations from the figures. We are planning a series of experiments using partly pre-programmed suggestions to the subjects and partly a completely free assessment of the AHP hierarchy of the measuring arguments.

A computer program may, moreover, include a check against too great an *intransitivity* of the individual weightings. If the evaluative preferences ($>_e$) are for example assessed as follows

$$(1)\ OP >_e 2xEP;\ (2)\ OP >_e 3xNG;\ (3)\ NG >_e 2xEP$$

the program may warn us that this is inconsistent, because the comparison between OP, EP and NG indicates that EP should be preferred to NG, provided that their evaluative relationship with the common point of comparison OP has been duly assessed.

6. THE CHOICE OF STRATEGY, SHOULD THE $P_{EV/MIN}$-VALUES NOT BE MET

The method presupposes that the evidential part of the decision is *either right or wrong*, although there may be more than two decision alternatives. But the distribution of possible *uncertainty* calls for closer analysis.

The starting-point of this method—as used in a simple form with the assumption of unitary disutility over different decision alternatives—was that

$$P_{EV/min/ET} > P_{EV/ET} \rightarrow \sim ET$$

following the axiom that the probability for $\sim$ET was $1 - P_{EV/ET}$. However, we have not accepted this axiom, and for this reason it may well be that $P_{EV/ET} < 1 - P_{EV/min/ET}$. Here the rational minimum certainty about the ET is not met. But if neither $P_{EV/min/ET}$ nor its counterpart for the ET, $1 - P_{EV/min/ET}$ are met, due to the twilight zone of evidential uncertainty, the linear strategy of Kaplan's formula does not work. How should one proceed in such a situation? There are two theoretical positions.

(I) The traditional approach of the burden-of-proof theory. Attention is focused only on the main evidence for the ET. If there is no sufficient certainty about the ET, it does not matter what the concrete EV of the evidence for the $\sim$ET is. The traditional idea of the burden of proof is that insufficient evidence for the ET automatically entails the $\sim$ET being taken as the factual premise of the decision—although at least adherents of the evidentiary value method admit that this normatively-based inference says nothing about the concrete evidence for the $\sim$ET.

(II) Alternatively, one may make certain *comparisons*:

(II:1) The simplest form is represented by the preponderance-of-evidence doctrine. Here, instead of resorting to traditional burden-of-proof rules, one compares the concrete EVs for the ET and the $\sim$ET and chooses the factual premise(s) of the decision on grounds of preponderance.

(II:2) A more sophisticated comparison is one between the P_{min}-va-

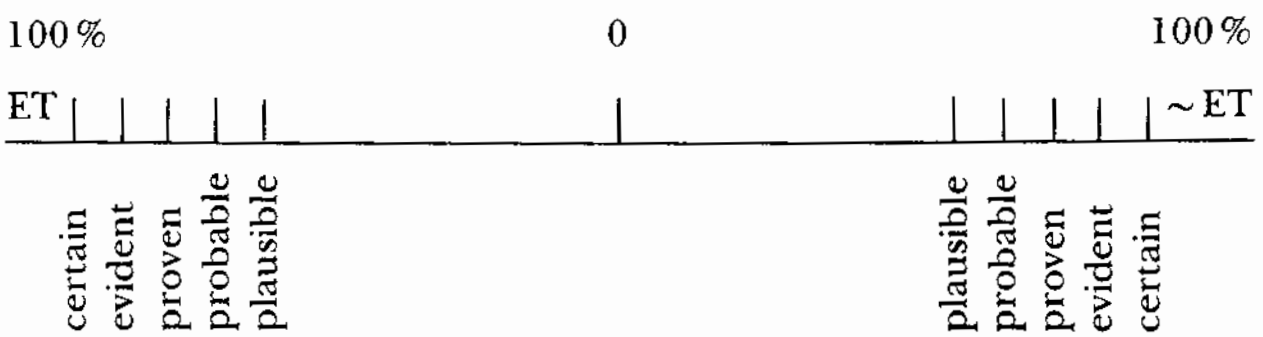
Our point is here that the *parties' awareness* of the criteria used when assessing the attitude towards uncertainty may have a function similar to that of pre-fixed criteria about sufficient evidence and the burden of proof.

One should not overestimate the impact of fixed rules upon the pre-trial and trial conduct of the parties. But this is an assumption pertaining to legal sociology, and we have no empirical evidence for our suggestion. In any case, the effect of awareness of guidelines probably differs little from the impact of fixed rules on sufficient evidence and the burden of proof.

We have already stressed that the method presented in this essay allows for *different strategies* as to evidential uncertainty: the one-sided strategy behind the traditional burden-of-proof rules may not be wholly justified. On the other hand, a straightforward comparison between the concrete EVs and the disutility-based $P_{EV/min}$-values may also be too schematic. We are trying, however, to develop "our" method in the light of empirical tests, and have chosen to suggest here answers to certain questions rather than attempting to present a set of conclusive solutions to all the problems involved.

7. CONCLUDING REMARKS

The assessment of evidentiary values and the minimum requirements for the sufficiency of evidence call for numerical operations. Lawyers are not accustomed to transforming their reasoning to figures, and it may be objected that numerical EVs or disutilities are too difficult to handle. But it is, rather, too difficult to use the expressions of ordinary language, due to their vagueness. Our empirical survey revealed an astonishing diversity here: on Ekelöf's scale of expressions used in Swedish legal language for indicating the degree of sufficiency of evidence the expression "probable" could be put between 1 and 90 %![34]



[34] Ekelöf in *Festschrift für Fritz Baur*, p. 352.

The placing of the words on the scale represents only rough estimates of the EVs required.

But despite Ekelöf's clarifying recommendations, the terminology of Swedish legislation was understood in a very heterogeneous manner. The situation in Finland was not different. Can the legislator exert any significant influence upon the sufficiency of evidence, if the language and thinking of the judges are as unclear as this? Thus, the use of figures—despite the difficulties involved—would ultimately improve the conditions of intersubjective rational discussion. Legislatory interventions would benefit, too, where intervention is considered advisable. For the method suggested here may also be used for general discussion about sufficient evidence and the burden of proof; the emerging criteria may then be compared with concrete situations where there may be grounds for revising the general criteria. One advantage of the method is that the structural unity of the framework allows comparison between abstract, typical situations and concrete cases.

Moreover, general rules may be established analytically, assessing the $P_{EV/min}$-criteria separately for the typical cases that would fall within the scope of the planned rule and then comparing their (perhaps weighted) values from the point of view of generalization: abstract rules are almost always in a certain sense compromises.

Although the method is complicated, we really believe in its future applicability. The calculations presuppose the use of computers and programs. But in many other decision-making situations access to computers has improved the rationality of hitherto very rough and goal-oriented assessments and evaluations of cost, prices, risks, etc.[35] Formal methods should be empirically established improvements of actual decision-making, i.e. rational and consistent models of actual reasoning. We think the method presented in this essay is such an improvement: the main ideas are based on our empirical studies of the ways Swedish and Finnish lawyers think about evidence, and on the argumentation used in discussion of sufficient evidence and the burden of proof.

We think, incidentally, that a somewhat similar assessment of "evidence" criteria could apply to problems of juristic interpretation, too. For (un)certainty about truth—interpreted truth—in the light of rules of experience and the disutility of a wrong decision is structurally similar to

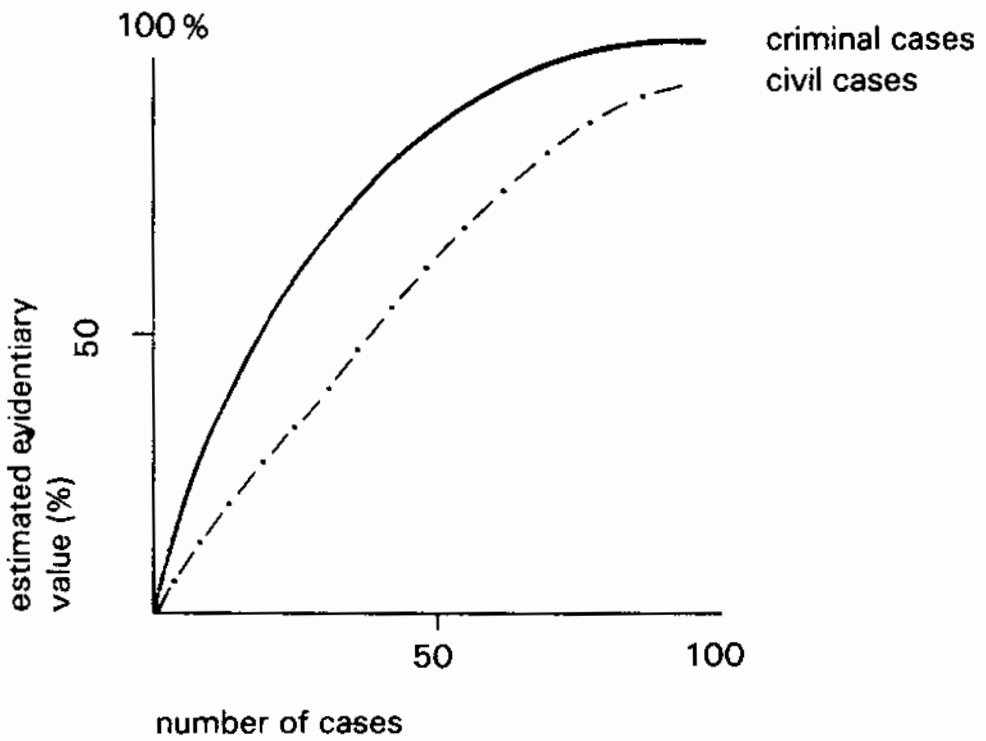
[35] One may refer to the application of decision-making theory to the strategies and tactics of games, voting, business, warfare, economic and social policy, etc. The theory provides models for assessing the optimal decision, even if certain decision parameters should be "fuzzy".

(un)certainty about the correctness of the interpretation of rules: the disutilities may also be assessed with the help of different goals and their weighted hierarchy. But it is of course impossible here to go into details of the problems of questions of law.

The elated vision of John Hus at the stake, already surrounded by flames: "Pravda viteži!"—the truth will win—will probably never be realized. But we believe that rationality will in the long run prevail, both in human decision-making in general and in judicial decision-making where uncertainty about facts can never be wholly avoided. We hope that the method presented here could at least promote discussion about making rational decisions under uncertainty.

The effect of the criteria for sufficient evidence depends on something unknown, namely the incidence of evidentiary values—and truth. That one is ready to accept a certain risk of wrong decisions does not *per se* mean that a corresponding number of wrong decisions will actually be made. This, again, depends on the actual reliability of the available information and on the correct assessment of the evidentiary values, not only on the criteria for the sufficiency of evidence.

One is probably justified in assuming—approximately, of course—the following curves:



There are in Sweden and Finland only $\leqslant 5$ % complete acquittals; moreover, one may assume that the probability of a case being brought to court upon only very weak evidence is small. The curve of the criminal cases with EVs > 50 % will therefore probably rise rather rapid-

ly. The number of cases with almost full certainty is not small, either (routine cases with conclusive evidence including a confession). Putting the criterion of sufficient evidence at 60, 70 or 80 % may for this reason not greatly affect the acquittal ratio or the actual risk of wrong decisions.

In civil litigation there are probably many more cases with rather low EVs. On the other hand, clear cases may be settled out of court. In any event, in civil cases the impact of the criterion of sufficient evidence will probably be greater than in criminal cases.

Assumptions about graphs and the truth behind them may influence the choice of strategy for cases where the P/min-values are not met either for the ET or for the ~ET. But to reason in e.g. the following manner would be rather dangerous: because false prosecutions are unusual, it will be rational to convict a second offender upon rather weak evidence, the only evidence for him being *not guilty* being his sturdy denial. Today the number of criminal cases with weak evidence is low because prosecutors know that strong evidence is required for conviction. For this reason they refrain from prosecuting, due to the disutilities involved when criminal trials result in acquittals, whether wrongful or not.[36]

But what is the actual risk of wrong decisions? In our poll study the guesses of Swedish judges varied much: several subjects stressed that one will never know the number of evidentiary mistakes.

[36] The disutility method may also be applied to prosecutorial decision-making. A wrong decision involving no prosecution is equivalent to a wrongful acquittal by the court. The disutility of a wrong decision against an innocent person is roughly:

the likelihood of conviction x the disutility of a wrong conviction.

One may, of course, assess the disutility of wrongful prosecution in a different and more specific manner.