

# EVIDENCE THEORY: A MATHEMATICAL FRAMEWORK FOR UNPREDICTABLE HYPOTHESES

*Guido Fioretti \**  
*Università di Siena*  
*Centro Sistemi Complessi*

## **Abstract #**

Shafer's evidence theory is a branch of the mathematics of uncertain reasoning that allows for novel possibilities to be conceived by a decision-maker. Many of its findings exhibits striking similarities with an alternative decision theory purported by Shackle in the 1950s, before expected utility maximisation monopolised the scene. Evidence theory does not attempt to formalise the emergence of novelties, but it is a suitable framework for reconstructing the formation of beliefs when novelties appear. An application to decision-making in the biotech and the pharmaceutical industry illustrates the potentialities of evidence theory, as well as its shortcomings.

**Keywords:** Shackle, Shafer, Cognition, Innovation, Biotechnology

---

\* Guido Fioretti, via Angeloni 11, I – 05100 Terni. E-mail: [fioretti@csc.unisi.it](mailto:fioretti@csc.unisi.it)

# I wish to thank Victoria Chick, Jean Czerlinski, Marco Dardi, Vittorio Emanuele Ferrante, Thorbjørn Knudsen, Klaus Nielsen, Bart Nooteboom and Germano Resconi for helpful comments and remarks. Furthermore, I feel indebted with Gail James for having produced the study that enabled me to sketch an empirical application of evidence theory.

## 1. Creative Decision-Making

How do people make their decisions when they conceive novel possibilities, and how do they evaluate innovative decisions that others made? How do managers decide to invest in a novel technology, and how do financial market analysts and credit institutions evaluate them? *Ex post*, given any set of possibilities decision-making can be modelled as utility maximisation. But *ex ante*, can we say anything about the process by which a possibility set is conceived in the mind of a decision-maker?

In the early days when expected utility maximisation had not yet imposed itself, G.L.S. Shackle was exploring the possibility of adopting a decision theory that, disregarding the emerging consensus on subjective probability theory (Ramsey 1931; De Finetti 1931; Savage 1954), would account for creativity and inventiveness in decision-making:

(...) we think of uncertainty as more than the existence in the decision-maker's mind of plural and rival (mutually exclusive) hypotheses amongst which he has insufficient epistemic grounds of choice. Decision, as we mean the word, is creative and is able to be so through the freedom which uncertainty gives for the creation of *unpredictable hypotheses*. Decision is not choice amongst the delimited and prescribed moves in a game with fixed rules and a known list of possible outcomes of any move or sequence of moves. There is no assurance that any one can in advance say what set of hypotheses a decision maker will entertain concerning any specified act available to him. Decision is thought and not merely determinate response.

(Shackle 1961, p. 6)

Shackle did not limit himself to wishful thinking. He laid down principles for an alternative decision theory that still deserve consideration.

Nonetheless, economics largely forgot Shackle. Lack of formalisation, of psychological foundations and of empirical applications are, surely, among the causes of this neglect. But independently of what Shackle might have done better, economics was simply taking another track. The booming economies of the 1960s were based on predictable events

and routine decisions; later on, the problems of stagflation were tackled by means of rational expectations over probability distributions of well-defined events, such as inflation levels.

Today, a turbulent economic environment and an accelerating rate of technological progress might bring Shackle back to the stage. But apart from relevance, the main reason for writing this essay is that, after Shackle had been forgotten, a suitable framework for the concepts that he had envisaged was made available by a mathematician, Glenn Shafer (1976), and that this formalisation remained unnoticed among economists hitherto. The converse is not true, since Shackle's writings received considerable attention by Shafer (Shafer 1976, p. 223). However, Shafer never bothered about propagating his ideas beyond a restricted circle of mathematicians.

Obviously, Shafer did not formalise creativity. However, he placed uncertain beliefs in a framework that can adapt to creativity.

Creativity means conceiving novel mental categories, containers that classify empirical evidence in a novel way and, possibly, are better able to discriminate the relevant aspects of emergent phenomena. Novel categories detect novel empirical evidence, that may be fragmentary, irrelevant, contradictory or supportive of particular hypotheses. Shafer did not attempt to say how novel evidence comes to be recognised, but provided rules for combining it with older ones. By doing this, he defined magnitudes that are surprisingly similar to Shackle's "potential surprise".

Since its first appearance in 1976 evidence theory is gaining increasing acceptance in the field of artificial intelligence, particularly in the design of expert systems (Shafer 1990, 1992). On the contrary, its popularity among probabilists remains low, since it requires epistemological assumptions that are at odds with those underlying probability theories.

The aim of this essay is to explain the foundations of evidence theory, establishing a link with Shackle's intuitions and applying it to a case-study in the biotech industry. Section

2 illustrates the basic concepts of evidence theory. Section 3 compares Shafer's and Shackle's uncertainty theories. Section 4 makes use of evidence theory in order to understand decision-making in the biotech industry. Finally, section 5 concludes.

## **2. The Frame of Discernment**

Subjective probability theory assumes that individuals are always able to conceive compound events out of union, intersection and complementation of a given list of elementary events. Novel events can be conceived, for instance out of the free space eventually left by a residual event. However, any novelty must derive from combination of given elementary possibilities by means of given rules. Implicitly, subjective probability theory adopts an algorithmic view of intelligence. Apart from the fact that this view is no longer tenable, it is obviously at odds with Shackle's idea of "unpredictable hypotheses".

On the contrary, evidence theory does not make any assumption concerning the way human imagination works. Simply, it describes decision-makers receiving information from different sources and evaluating to what extent the evidence that they provide is compatible or contradictory. Once conflicts have been evaluated, a decision-maker may hold beliefs on whatever possibility he may envisage. The role of evidence theory is that of telling the decision-maker which evidence supports the possibility that he is considering.

Clearly, the idea that decision-makers may not draw all logical inferences is one of bounded rationality (Simon 1982). However, the idea that not all inferences can be produced

by means of logical combination of given items goes far beyond that, since it ultimately relies on Gödel’s impossibility theorem. Evidence theory includes probability theory as a special case that obtains – among else – under the hypothesis that decision-makers are able to consider any combination of a given set of events and that no other event can be conceived.

Figure 1 illustrates this concept on a very simple example, namely a decision-maker who is envisaging two possibilities A and B in a possibility set  $\Theta$ . According to probability theory (left), the complementary of A is the whole dashed area of  $\Theta$ . In fact, probability theory assumes that this decision-maker is able to conceive all possibilities in  $\Theta$  – to be honest, the very idea of envisaging only A and B would not make much sense within probability theory. On the contrary, according to evidence theory (right) the complementary of A is just B. In fact, the set of possibilities is not given once and for all, but it may change with time in unpredictable ways; at this moment, it only entails A and B. As we shall see in the next section, this difference is important in order to understand the formalisation of Shackle’s idea of “potential surprise”.

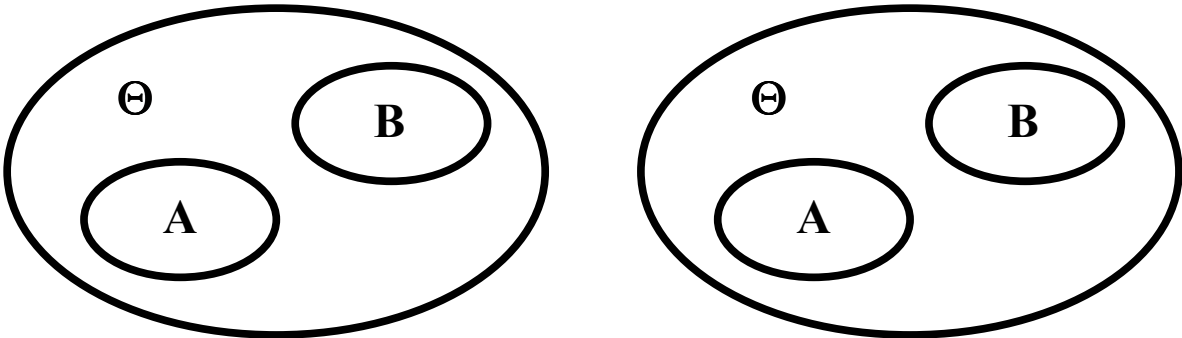


Figure 1

The dashed area depicts the complementary set of A according to probability theory (left) and according to evidence theory (right).

According to evidence theory, possibility sets are mental representations of empirical evidence in an individual's mind. In order to stress their cognitive nature, Shafer preferred to use another term:

It should not be thought that the "possibilities" that comprise [a set]  $\Theta$  will be determined and meaningful independently of our knowledge. Quite to the contrary:  $\Theta$  will acquire its meaning from what we know or think we know; the distinctions that it embodies will be embedded within the matrix of our language and its associated conceptual structures and will depend on those structures for whatever accuracy and meaningfulness they possess. In order to emphasize this epistemic nature of the set of possibilities  $\Theta$ , I will call it the *frame of discernment*. When a proposition corresponds to a subset of a frame of discernment, I will say that the frame *discerns* that proposition.

(Shafer 1976, p. 36)

What is left for mathematics, if frames of discernment evolve in unpredictable ways and depict unpredictable hypotheses? Evidence theory consists of a set of tools for describing how empirical evidence is combined on a given frame of discernment. Eventually, contradictory evidence may trigger efforts to develop a new frame, so although mathematics cannot tell which unpredictable hypotheses will emerge (otherwise, they would be predictable!), it can tell whether the available frame provides a coherent interpretation of reality.

Empirical evidence is assumed to be available as sets of numbers  $\{m(A_1), m(A_2), \dots, m(\Theta)\}$  which represent the amounts of evidence that support subsets  $\{A_1, A_2, \dots\}$  of a frame of discernment  $\Theta$ , respectively. Each set of numbers  $\{m_1, m_2, \dots, m_\Theta\}$  is called a *body of evidence*.

Numbers  $\{m_1, m_2, \dots\}$  represent amounts of empirical evidence supporting alternative possibilities  $\{A_1, A_2, \dots\}$ . Subsets  $A_i$  may not be disjoint.

Although not essential for the theory, numbers  $m$  are generally normalised to satisfy

$$\sum_i m(A_i) + m(\Theta) = 1 \quad (1)$$

where  $m(\Theta) > 0$  eventually means that a decision-makers thinks that something may happen, which he is not even able to imagine<sup>1</sup>. Note that this framework is profoundly different from that of sub-additive probabilities (Koopman 1940a, 1940b, 1941; Choquet 1953-54; Good 1952, 1962; Smith 1961, 1965; Williams 1976; Gilboa 1987; Schmeidler 1989; Sarin and Wakker 1992, 1994). In fact, although sub-additive probabilities may fail to sum up to unity because of lack of information regarding a given set of possibilities, the rest of the probability mass cannot be left at disposal of “unpredictable hypotheses”.

In general, empirical evidence may be provided by different, independent sources. Let us suppose that a decision-maker has two sources of information, which provide bodies of evidence  $\{m(A_1), m(A_2), \dots, m(\Theta)\}$  and  $\{m(B_1), m(B_2), \dots, m(\Theta)\}$ , respectively. Subsets  $A_i$  and  $B_j$  may represent conflicting evidence, in which case they are disjoint sets, or they may represent (partially) coherent evidence, in which case they (partially) overlap.

Dempster-Shafer rule enables us to compute a new set of numbers  $m$  that represent combined evidence (Dempster 1966, 1967, 1968; Shafer 1976). According to Dempster-Shafer rule, for any possibility  $C$  the combined evidence provided by the above sources is:

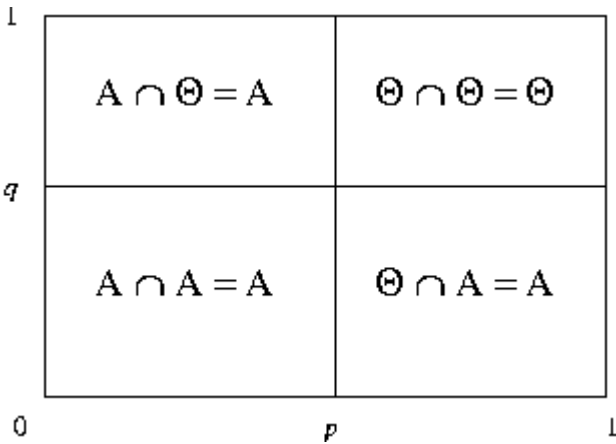
$$m(C) = \frac{\sum_{A_i \cap B_j = C} m(A_i) m(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m(A_i) m(B_j)} \quad (2)$$

where  $\emptyset$  represents the void set. Note that the Dempster-Shafer rule is independent of the order in which bodies of evidence arrive.

The numerator of Dempster-Shafer rule (2) represents the influence of aspects of the second body of evidence that confirm the first one. The denominator represents the influence of aspects of the second evidence that contradict the first one. Dempster-Shafer rule says that compatible evidence on a possibility C must be evaluated as a fraction of total compatible evidence.

The meaning of Dempster-Shafer rule (2) can be easily illustrated in the simple case of two bodies of evidence focused on one single possibility A (Shafer 1976, p.76). Suppose that one body of evidence is  $\{m(A)=p, m(\Theta)=1-p\}$  and that another body of evidence is  $\{m(A)=q, m(\Theta)=1-q\}$ . Figure 2 illustrates the ensuing frame of discernment.

In this simple case, only the evidence on  $\Theta$  does not support A. Thus, the second term in the denominator of (2) is  $m(\Theta)m(\Theta) = (1-p)(1-q)$ . Consequently, the total evidence in favour of A – the denominator of (2) – is  $1-(1-p)(1-q)$ . Out of this amount, only the fraction  $pq/1-(1-p)(1-q)$  is supported by both bodies of evidence.




---

<sup>1</sup> Mental habits stemming from probability theory led Shafer to call numbers  $m$  "basic probability assignments". This name is misleading. In fact, although numbers  $m$  are required to sum up to unity, they are not probabilities because they are defined on a set - the frame of discernment - which is not an algebra.

Figure 2

A frame where only one possibility is discerned,  $A$ , by two bodies of evidence placed on orthogonal axes. The upper right square must be excluded from computation of the hypothesis  $H \equiv A$ .

Dempster-Shafer rule (2) does not apply if two bodies of evidence are completely contradictory. It only makes sense if  $\sum_{A_i \cap B_j = \emptyset} m(A_i)m(B_j) < 1$ . If two bodies of evidence are completely contradictory, they can be handled as one single body of evidence over alternative possibilities whose numbers  $m$  must be re-scaled in order to comply with (1).

Finally, there exists a very special case where Dempster-Shafer rule (2) boils down to Bayes's rule of conditioning. Let us suppose that the first body of evidence is  $\{m(A_i)\}$  and that the second body of evidence is  $\{m'(B)=1\}$ . Let us apply Dempster-Shafer rule in order to find what evidence is left on a specific possibility  $C$  once  $\{m'(B)\}$  arrived. The numerator of (2) becomes  $\sum_{A_i \cap B=C} m(A_i)m'(B) = \sum_{A_i \cap B=C} m(A_i) = m(C \cap B)$ . The denominator becomes  $1 - \sum_{A_i \cap B=\emptyset} m(A_i)m'(B) = 1 - \sum_{A_i \cap B=\emptyset} m(A_i) = 1 - m(\bar{B}) = m(B)$ . Thus, Dempster-Shafer rule can be written as  $m(C|B) = m(C \cap B)/m(B)$ .

However, in spite of formal connections evidence theory refers to a setting that is quite different from the one that inspires probability theory (Shafer 1981, 1982). In fact, probability theory understands uncertainty in terms of similarity of the situation at hand with a game of chance. On the contrary, evidence theory understands uncertainty in terms of similarity of the situation at hand with the evaluation of testimonies in a trial (Shafer 1986b). Testimonies may partially confirm or contradict one another. The job of evidence theory is that of yielding a coherent picture for the judge to take stance.

Originated by George Hooper in 1699 and improved by Johann Heinrich Lambert in 1764, this trial-based approach was obscured by Laplace in the 19th century and re-discovered in the early 1960s by a Swedish professor of law, Per Olof Ekelöf (Shafer 1986b). Between the late 1960s and early 1970s, Arthur Dempster and Glenn Shafer provided a formalisation of the general case (Dempster 1966, 1967, 1968; Shafer 1976).

Trials and games of chance are quite different settings. Since these canonical situations suit particular decision problems to a greater or lesser extent, it may be wise to choose among them according to the situation to be described (Shafer and Tversky 1985). Possibly, a CEO who is listening to experts presenting scenarios in order to help making a crucial investment decision, or an analyst of a credit institution who is evaluating reports on a prospective borrower, is better approximated by a judge than by a gambler.

Ultimately, evidence theory aims at evaluating the confidence of a judge in the truth of the possibilities that he envisages. This evaluation is summarised by a *belief function*.

Belief functions express beliefs in the hypotheses entertained by a judge. The idea underlying belief functions is that hypothesis  $H$  is supported by any evidence  $A_i$  that refers to  $H$ , or aspects of  $H$ , but has no relationships with anything that is out of the scope of  $H$ .

Suppose that the evidence provided by different testimonies has been combined in order to yield one single, comprehensive body of evidence. Given a frame of discernment  $\Theta$  and a body of empirical evidence  $\{m(A_1), m(A_2), \dots, m(\Theta)\}$ , the belief committed to hypothesis  $H \subset \Theta$  is:

$$Bel(H) = \sum_{A_i \subset H} m(A_i) \quad (3)$$

Condition (1) implies that  $Bel(\Theta) = 1$ . However, the belief committed to  $\Theta$  cannot be distributed to its subsets.

Expression (3) means that all empirical evidence that supports hypothesis H, and nothing else, contributes to the belief that H is true. On the contrary, empirical evidence that supports H *and* some other part of the frame of discernment at the same time, cannot be taken to support H.

### 3. Shackle and Shafer

Let us examine the differences and similarities between Shackle and Shafer. Most of them had been already noted by Shafer himself (Shafer 1976, p. 223), although cursorily.

Shackle reasoned that, in order to leave room to unpredictable hypotheses, an uncertainty variable should not distribute a unit mass on a given list of possibilities, as probability does (Shackle 1961, p. 47). According to Shackle, whenever a probabilistic model needs to contemplate a residual event, a non-distributional uncertainty variable would be in order (Shackle 1961, p. 49).

Evidence theory fulfils this requirement. In fact, if an excess mass  $m(\Theta) > 0$  is available, this is not distributed to the possibilities that are envisaged at that time. On the contrary,  $m(\Theta)$  is kept aside for unpredictable hypotheses that may eventually materialise.

Once again, it is in order to stress that sub-additive probabilities keep aside the complement to one in order to distribute it to themselves when better information is available.

Shackle thought of a non-distributional uncertainty variable that should have been called "potential surprise":

It is the degree of surprise to which we expose ourselves, when we examine an imagined happening as to its possibility, in general or in the prevailing circumstances, and assess the obstacles, tensions and difficulties which arise in our minds when we try to imagine it occurring, that provides the indicator of degree of possibility. This is the surprise we *should* feel, if the given thing *did* happen; it is *potential* surprise.

(Shackle 1961, p. 68)

Shafer identified Shackle's "potential surprise" with  $Bel(\bar{H})$ , which measures the surprise that one would feel if a hypothesis entailed in  $\bar{H}$  would occur (Shafer 1976, p. 225). Remember that, contrary to probability theory,  $\bar{H}$  does not denote *all* possible alternatives to H that can be constructed by applying given rules to elementary events, but only those that a decision-maker has been able to imagine, with all the "obstacles, tensions and difficulties" that this process requires. Figure 1 illustrates the difference.

Why not  $Bel(H)$ , which is intuitively more immediate? Shackle made it clear that it was a stratagem aimed at avoiding that direct examination of one set of possibilities would distract the modeller from all other possibilities that could be thought. Shackle feared that examination of a set of possibilities would be a Trojan horse for a distributional uncertainty variable to creep in:

To invert a problem is a well-known resource of the mathematician. We too, in adopting as our uncertainty-variable a measure of *disbelief*, are in a way inverting our problem, and we are thereby guaranteeing our solution against any attempt to turn it back into a distributional variable. Formally, it would be open to us to define a measure of possibility directly. A zero value of this measure would stand for impossibility, an absolute maximum value for perfect possibility. But it would then be tempting for those who favour a distributional solution to

add together (contrary to reason and to the nature of the problem) the respective degrees of possibility assigned to the members of an exhaustive set of hypotheses, and to treat the resulting total as unity. This total could then, at no greater cost in perversity, be looked upon as 'distributed' over the various rival hypotheses. By contrast, when we represent perfect possibility by a zero value of a measure of disbelief, no such summation can in general be meaningfully carried out, since, for example, the sum of a set of zero values is zero.

(Shackle 1961, p. 75)

Notwithstanding Shackle's own choice, several writers who tried to build on his intuitions turned to non-distributional variables that express belief rather than disbelief. Essentially, Katzner's "potential confirmation" (Katzner 1986) or Ford's "degree of credibility" (Ford 1987) can be understood as  $Bel(H)$ .

However, both Shackle and Shafer suggested that both a measure of "belief" and a measure of "disbelief" should be used at the same time. In fact, Shackle wrote that an individual's belief should be expressed by the potential surprise of a hypothesis together with the potential surprise of its contradictory:

When we consider together as one whole all the hypotheses which are rivals to a given hypothesis H, they constitute the *contradictory* of H. If, then, we wish to express the individual's *degree of belief* in a hypothesis, we shall think of this degree of belief as consisting in a degree of potential surprise associated with the hypothesis, and in another degree associated with its contradictory.

(Shackle 1961, p. 73)

Quite similarly, Shafer wrote:

One's beliefs about a proposition H are not fully described by one's degree of belief  $Bel(H)$ , for  $Bel(H)$  does not reveal to what extent one doubts  $\bar{H}$  - i.e., to what extent one believes its negation  $\bar{H}$ . A fuller description consists of the degree of belief  $Bel(H)$  together with the *degree of doubt*  $Dou(H) = Bel(\bar{H})$ .<sup>2</sup>

(Shafer 1976, p. 43)

Thus, it appears that both Shackle and Shafer suggested to use both  $Bel(H)$  and  $Bel(\bar{H})$ . Shackle started from  $Bel(\bar{H})$  and added  $Bel(H)$ , whereas Shafer started from  $Bel(H)$  and added  $Bel(\bar{H})$ , but this is not a big difference.

Combination rules are another interesting issue to compare. Shafer himself noted that the rule suggested by Shackle in order to deal with rival hypotheses coincide with a special case of (3), namely the one that obtains when evidence is not contradictory. Compare the following statements by Shackle with a theorem that Shafer obtained for what he called “consonant belief functions”:

We would say that an individual's degree of belief in a hypothesis can be easily and exactly expressed by means of the potential surprise he assigns to the least (potentially) surprising rival hypothesis.

(Shackle 1961, p. 71)

At least one member of an exhaustive set of rival hypotheses must carry zero potential surprise.\*

\* But it is possible for all the rival hypotheses which are in any degree particularized or specified to carry potential surprise greater than zero, only the residual hypothesis carrying zero potential surprise.

(Shackle 1961, p. 81)

Suppose  $\Theta$  is a finite set. Then a function  $Bel: 2^\Theta \rightarrow [0,1]$  is a consonant belief function if and only if  $Bel(\Theta)=1$  and the function  $Dou: 2^\Theta \rightarrow [0,1]$  given by  $Dou(H) = Bel(\bar{H})$  obeys these two rules:

(1)  $Dou(H) = \min_{\theta \in H} Dou(\{\theta\})$  for all non-empty  $H \subset \Theta$ .

(2) There exists a  $\theta \in \Theta$  with  $Dou(\{\theta\}) = 0$ .

(Shafer 1976, p. 224)

Figure 3 illustrates the meaning of the above claims in the more intuitive case of  $Bel(H)$ . Let  $\{m(A_1), m(A_2)\}$  be a body of evidence that supports a belief on H. If this

---

<sup>2</sup> Shafer actually used the symbol A where I wrote H. I did so in order to stress similarities with Shackle's

evidence is not (even partially) contradictory and if we abstract from the trivial cases where  $H$  is either included or includes both  $A_1$  and  $A_2$ , it must be either  $A_1 \subset H \subset A_2$  or  $A_1 \supset H \supset A_2$ . Thus, according to (3) it is either  $Bel(H) = m(A_1)$  or  $Bel(H) = m(A_2)$ .

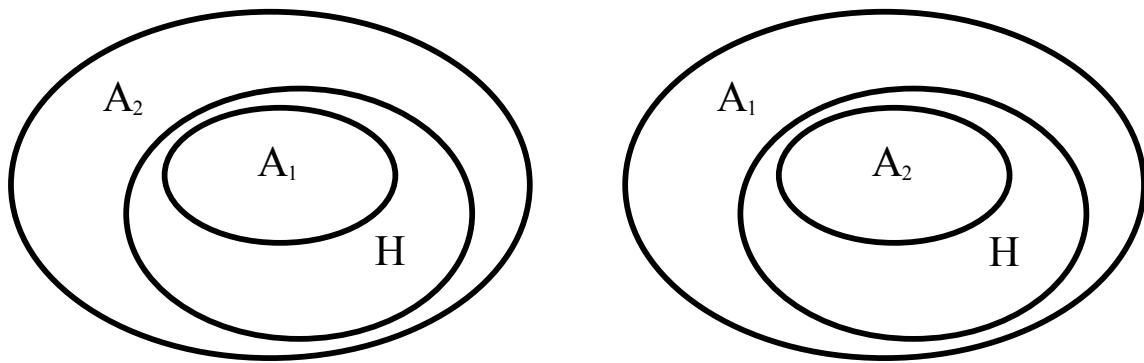


Figure 3

If evidence is not contradictory, not even partially contradictory, then belief coincides with the least comprehensive testimony.

The ability of evidence theory to deal with contradictory evidence by means of Dempster-Shafer rule (2) is an important generalisation of Shackle's intuitions. A few attempts at an empirical test of Shackle's ideas support the usefulness of the concept of "potential surprise" at least in *some* decision settings (Hey 1985; Ford and Ghose 1994a, 1995a, 1995c), but discard Shackle's suggestions for calculating beliefs (Ford and Ghose 1994b, 1995b).

Thus, Shackle actually covered a subsection of evidence theory. This part of evidence theory is now called *possibility theory* by mathematicians (Dubois and Prade 1988), since in this case belief functions can be understood as possibility measures. Evidence theory can be linked to fuzzy sets theory (Klir and Yuan 1995), and both of them can be seen as deriving from modal logic (Resconi, Klir and St.Clair 1992; Klir, St.Clair and

---

"hypotheses".

Harmanec 1993; Harmanec, Klir and Resconi 1994; Klir and Harmanec 1994; Resconi, Klir, Harmanec and St.Clair 1995).

This link with fuzzy sets theory made evidence theory even closer to Shackle's intuitions. In fact, once belief functions are interpreted as measuring the degree of membership of H to a fuzzy frame  $\Theta$  in terms of degrees of possibility, with  $Bel(H) = 0$  if H is impossible,  $Bel(H) = 1$  if H is perfectly possible and  $Bel(H) \in (0,1)$  if it is in between, some seemingly obscure statements made by Shackle can be readily understood:

We can go far in the analysis of decision by means of a mere possible-impossible dichotomy. Uncertainty as we ordinarily understand this word, however, includes a more subtle meaning of 'doubt' than the mere plurality of hypotheses placed in the 'possible' as against the 'impossible' class. There seems in the observed or reported working of our minds to be a faculty of adjudging degrees of possibility (...). (...) It is partly in the formation of such judgements that we think of the essentially new as being able to enter. It is here that there can be inspiration. This is the locus of creative decision.

(Shackle 1961, p. 12)

Shackle stressed that decision-making is a creative process. Creativity, within the framework of evidence theory, reflects in changing one's frame of discernment. Even if it is impossible to predict which "unpredictable hypotheses" will be discerned in the future, Shafer attempted to identify the forces that make a frame expand or contract:

Like any creative act, the act of constructing a frame of discernment does not lend itself to thorough analysis. But we can pick out two considerations that influence it: (1) we want our evidence to interact in an interesting way, and (2) we do not want it to exhibit too much internal conflict.

(...) Since it depends on what we are interested in, any judgment as to whether our frame is successful in making our evidence interact in an interesting way is a subjective one. But since interesting interactions can always be destroyed by loosening relevant assumptions and thus enlarging our frame, it is clear that our desire for interesting interaction will incline us towards abridging or tightening our frame.

Our desire to avoid excessive internal conflict in our evidence will have precisely the opposite effect: it will incline us towards enlarging or loosening our frame. For internal

conflict is itself a form of interaction - the most extreme form of it. And it too tends to increase as the frame is tightened, decrease as it is loosened.

(Shafer 1976, p. 280)

Abridging, tightening, contracting a frame of discernment amounts to excluding some possibilities in order to focus on more interesting interactions. Likewise, enlarging, loosening or expanding a frame of discernment amounts to thinking of new possibilities in order to adjust contradictory empirical evidence. For instance, in the 1980s car manufacturers greatly increased the number of optionals: in terms of evidence theory, they expanded their frame of discernment in order to satisfy idiosyncratic needs that risked to generate contradictory specifications. On the contrary, in the 1920s Henry Ford contracted his frame of discernment in order to focus on the interesting fact that a large market for cheap, reliable automobiles was growing.

Shafer also warned that nothing warrants that individuals use only one single frame of discernment at a time:

Since we do find it necessary to adjust our assumptions to our circumstances, we use many different and incompatible frames of discernment in our practice of probable reasoning. In fact, we often consider many different frames more or less simultaneously. Even when our attention is sharply focused we tend to experiment to some extent at varying our assumptions, and as we shift the focus and level of generality of our attention we tend to vary these assumptions even more, sometimes going so far as to take for granted on one occasion something that might be the focus of our questioning on another. Consequently we will construct many different frames of discernment, which will vary in their success in organizing our experience and thus in their ultimate acceptance.

(Shafer 1976, p. 281)

This insight further complicates the picture of human reasoning and, coherently with the rest of evidence theory, it stresses that the mathematics of uncertain reasoning cannot be developed independently of psychology.

#### 4. Pharmaceuticals and Biotechnology

Biotechnology is an instance of a radical innovation that is questioning established oligopolies and traditional industry strategies. *Ex post*, we can say that for the pharmaceutical industry of a few decades ago biotechnology was a really unpredictable hypothesis. Thus, it is interesting to observe the evolution of the frames of discernment of decision-makers involved in biotech and pharmaceuticals and to check whether evidence theory is useful in order to understand on which grounds decisions have been taken.

Managers's views can be evinced by means of linguistic analysis of documents produced by their firms in order to explain their strategies; typically, letters to shareholders are examined (Tucker Wrightson 1976; Sigismund Huff, Narapareddy and Fletcher 1990). This technique allows to reconstruct which possibilities were envisaged by managers, as well as which causal links they expected to connected them to one another. Possibilities and causal links describe the strategies that managers ascribed to their competitors or chose for themselves and, taken together, constitute their *cognitive map* at the time the document was produced (Axelrod 1976a, 1976b; Weick 1990; Sigismund Huff 1990).

In the case of the emerging biotech industry and its relationships with the broader pharmaceuticals industry I shall draw on a study of documents produced by the *Biotechnology Industry Organization* and by *Ernst & Young*, a Wall Street analyst who is acquainted with both biotech and pharmaceutical industry (James 1996, 2000). Evaluation of empirical evidence takes place at three levels. Firstly, biotech firms evaluate empirical

evidence in order to draw inferences about the future of their industry and their relationships with pharmaceutical firms. Secondly, pharmaceutical firms evaluate empirical evidence in order to predict the evolution of their relationships with biotech firms. Thirdly, Wall Street evaluates the (partially) conflicting evidence presented by emerging biotech firms and incumbent pharmaceuticals.

Biotechnology is generally foreign to large pharmaceutical companies, so it offers a scope for innovative start-ups. However, large R&D expenditures and the need of a large organisation in order to market drugs among the public makes it impossible for start-ups to finance their activities by market sales. Furthermore, the long time needed to develop a new drug (8-10 years) makes it difficult for them to raise money in the financial market.

Pharmaceutical firms, in their turn, are interested in building-up expertise in biotechnology. Thus, from the very beginning biotech and pharmaceutical firms were forced to forge alliances. According to the contracts that they stipulated, biotech firms carried out research on behalf of pharmaceutical firms, which eventually exploited their scientific findings. However, these collaborations started with opposite views: on the one hand, biotech start-ups strove to become large, vertically integrated pharmaceutical firms; on the other hand, pharmaceutical firms aimed to select start-ups to acquire and develop biotech in-house.

In 1986, biotech firms were sure that sooner or later collaborations with pharmaceutical firms would come to an end and that they would become fully independent firms, drawing revenues from market sales. Biotech firms would either become vertically integrated pharmaceutical firms, or independent niche firms, or they would suffer competition by pharmaceutical firms if these would arrive at mastering biotechnology. They based their inferences on the following evidence: Biotechnology offers great profit

opportunities ( $B_1$ ); We are too small to market our products ( $B_2$ ); We cannot draw resources from financial markets ( $B_3$ ).

It seems sensible to assume that these three pieces of evidence are independent of one another, so they must be treated as three separate bodies of evidence. Let us assume that  $E'_B(1986) = \{m(B_1) = 0.99, m(\Theta) = 0.01\}$ ,  $E''_B(1986) = \{m(B_2) = 0.9, m(\Theta) = 0.1\}$ ,  $E'''_B(1986) = \{m(B_3) = 0.8, m(\Theta) = 0.2\}$  and that  $B_1, B_2, B_3$  are arranged as in figure 4a.

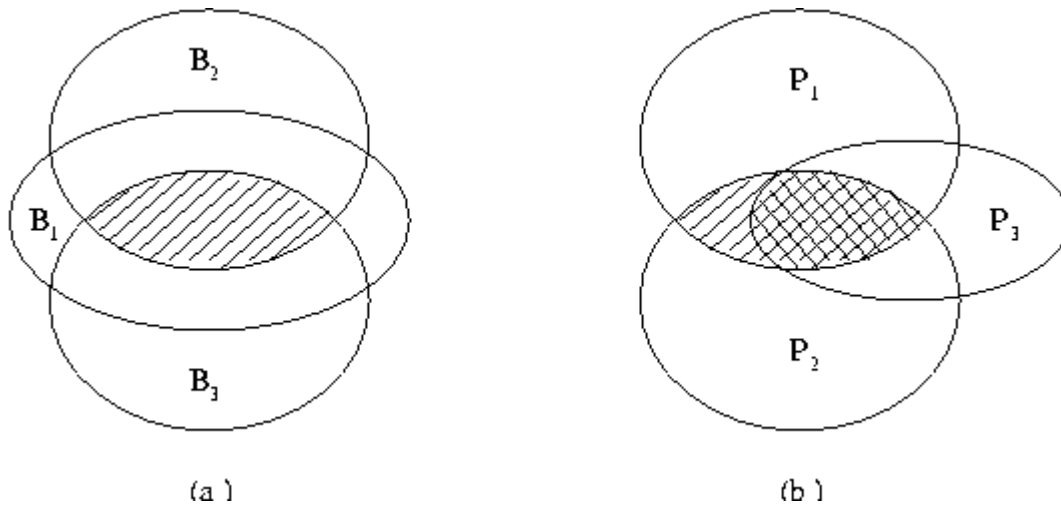


Figure 4

The frames of discernment of biotech firms (a) and pharmaceutical firms (b), 1986. Both actors relied on evidence that suggested to enter collaboration contracts, but pharmaceutical firms made use of additional evidence in order to claim that in the end they would acquire biotech start-ups.

Managers of biotech firms are interested in  $Bel(B_1 \cap B_2 \cap B_3)$ . In order to calculate this, they first apply (2) to the combination of  $E'_B(1986)$  and  $E''_B(1986)$ , yielding  $m(B_1 \cap B_2) = (0.99 \times 0.9) / (1 - 0.01 \times 0.1) \cong 0.892$ , and also  $m(B_1 \cap \Theta) = m(B_1) = (0.99 \times 0.9 + 0.99 \times 0.1) / 1 = 0.99$ ,  $m(\Theta \cap B_2) = m(B_2) = (0.99 \times 0.9 + 0.01 \times 0.9) / 1 = 0.9$  and  $m(\Theta) = (0.01 \times 0.1) / 1 = 0.001$ . Thus, the ensuing body of evidence is  $\{m(B_1 \cap B_2) \cong 0.892,$

$m(B_1 - B_1 \cap B_2) \cong 0.098$ ,  $m(B_2 - B_1 \cap B_2) \cong 0.008$ ,  $m(\Theta) = 0.001$ }. By means of a second application of (2) we can combine this body of evidence with  $E_B'''(1986)$ . This yields, among else,  $m(B_1 \cap B_2 \cap B_3) \cong (0.892 \times 0.8) / (1 - 0.001 \times 0.3) \cong 0.714$  and  $m(\Theta) = 0.0002$ . Since there is no other piece of evidence bearing on  $B_1 \cap B_2 \cap B_3$ , a trivial application of (3) yields  $Bel(B_1 \cap B_2 \cap B_3) \cong 0.714$ .

Thus, it appears that biotech firms had some good reasons to believe that  $B_1 \cap B_2 \cap B_3$  was true, which implies entering collaborations with pharmaceutical firms. However, this evidence does not support their confidence that one day they would become fully independent firms, able to draw revenues from market sales.

In the same year, pharmaceutical firms were absolutely sure that collaborations would come to an end because sooner or later they would either acquire biotech firms, or they would become able to develop biotechnology in-house. They based this conviction on the following evidence: Many pharmaceutical patents are expiring, so we need biotechnologies ( $P_1$ ); Biotech firms have valuable competences ( $P_2$ ); We have large financial resources ( $P_3$ ).

Again, it seems sensible to assume that these three pieces of evidence are independent of one another, so they must be treated as three separate bodies of evidence. Let us assume that  $E_p'(1986) = \{m(P_1) = 0.9, m(\Theta) = 0.1\}$ ,  $E_p''(1986) = \{m(P_2) = 0.9, m(\Theta) = 0.1\}$ ,  $E_p'''(1986) = \{m(P_3) = 1\}$  and that  $P_1, P_2, P_3$  are arranged as in figure 4b.

A first application of (2) and (3) yields  $Bel(P_1 \cap P_2) \cong 0.818$ . Since  $P_1 \cap P_2$  is sufficient to enter collaborations with biotech firms, pharmaceutical firms had good reasons to do it. However, knowing with certainty to have much larger financial resources than

biotech firms implies that  $Bel(P_1 \cap P_2 \cap P_3) \cong 0.818$  as well. Thus, their conviction that in the long run they would acquire biotech firms seems to rest on solid grounds.

How did market analysts judge the evidence provided by biotech and pharmaceutical firms? Considering that all the evidence provided by pharmaceutical firms is compatible with the evidence provided by biotech firms, iterated application of Dempster-Shafer rule yields  $Bel(B_1 \cap B_2 \cap B_3 \cap P_1 \cap P_2 \cap P_3) \cong 0.643$ . Thus, it appears that in 1986 market analysts could reasonably think that, in the long run, pharmaceutical firms would buy biotech start-ups. Examination of their cognitive map confirms that this is what they actually believed.

This finding seems to suggest that evidence theory can be easily applied to concrete situations and that it is a powerful tool in order to reconstruct how decision-makers develop their beliefs. However, in 1986 all the actors had little experience with biotech so their beliefs reflected immediate evidence very closely. In the subsequent years, beliefs relied on older evidence as well. Thus, application of evidence theory would much more problematic, simply because it would not be clear on which evidence judgements are made.

This is most evident in the 1990 “poison pill” episode. Reacting to a 1989 warning by a prominent market analyst, biotech firms were very much concerned that collaborations with pharmaceutical firms might end up with biotech firms giving away their expertise to their partners and consequently lose their market value. The 1990 cognitive map of biotech firms has none of the arguments that characterise their previous and subsequent maps, like e.g. their technical excellence, the difficulty of drawing revenues from market sales, and so on. Rather, it only considers evidence of collaborations where a biotech firm lost control over its technology and in the end it could not attract new partners. It is striking that only in 1990 biotech firms realised this possibility, and that they were so shocked that any other possible development was forgotten.

How did market analysts react? Simply, they ignored this 1990 biotech firms hysteria. The cognitive map of market analysts in 1990 is very much akin to that of 1989. This is a very important remark, because if we would apply evidence theory to the 1990 cognitive maps of biotech firms and pharmaceutical firms in order to reconstruct market analysts' conclusions, we would fail to provide a sensible account. Since market analysts relied on evidence collected along experiences with biotech that already spanned several years, evidence theory should be rather applied to *all* the evidence that they had collected at the time they provided their evaluation.

Another interesting issue is that, on some occasions, biotech and pharmaceutical firms evaluated evidence which market analysts had suggested to highlight. In 1990, market analysts suggested that collaborations between biotech and pharmaceutical firms could be beneficial for both partners. In fact, biotech firms would obtain funds and learn marketing and organisational skills; in their turn, pharmaceutical firms would obtain a permanent window on novel technological developments. They called it a "Win-Win" relationship.

In 1991, *both* biotech and pharmaceutical firms claimed that biotech firms, once acquired, lose their innovative capacity. Sure, this conviction stemmed out of experiential evidence. Nevertheless, analysts' concept of a "Win-Win" relationship might have helped to highlight it.

In any case, this piece of evidence marked the end of the suspicious collaboration model of the previous years and opened a possibility for true, lasting collaborations. Since 1991, all cognitive maps presented this argument. In terms of evidence theory, we would say that new  $B_4$ ,  $P_4$  arrived such that  $B_1 \cap B_2 \cap B_3 \cap B_4 \equiv P_1 \cap P_2 \cap P_3 \cap P_4$ .

Finally, it should be remarked that evidence theory is good at combining evidence about novel possibilities with evidence about old possibilities, but it is unable to say how

novel possibilities – Shackle’s “unpredictable hypotheses” – arise in the minds of decision-makers. Market analysts’s 1993 cognitive map is a clear instance. In fact, although the cognitive maps of biotech and pharmaceutical firms did not change through 1991, 1992 and 1993, in this last year market analysts provided a much more detailed framework for analysing inter-firm collaborations. This evolution was not triggered by any novel empirical evidence, but rather suggests which is the evidence decision-makers should look for. In Shafer’s own words, in 1993 market analysts chose to enlarge their frame of discernment.

## 5. A Few Concluding Remarks

In the example expounded in the previous section, numbers  $m$  had been simply guessed. This may be acceptable in an illustrative example, but it would not be acceptable in an empirical application.

How do we measure evidence? Although Shafer proposed a method based on counting instances, it is evident that the praxis of decision-making is influenced by other factors as well.

A first crucial issue is that of *reliability* of testimonies. Shouldn’t we weigh empirical evidence according to the reputation of their source? A second one is *independence* of testimonies. If testimonies are only partially independent, shouldn’t we weigh bodies of evidence according to the extent to which they are independent of one another?

Of course, we should. But how do we do that? It is clear that evidence theory, possibly to an even greater extent than probability theory, cannot be mechanically applied to real-world problems. If care is taken in order to ensure that all relevant evidence has been included and that it has been properly discounted for reliability and independence of testimonies, then evidence theory offers a sensible representation of the way beliefs are formed.

However, it should be remarked that evidence theory is only concerned with beliefs whose truth or falsity can be ascertained by facts. What about beliefs that are not supported by immediate empirical evidence, which nonetheless inspire crucial decisions? At the beginning of the previous section, we remarked that in 1986 biotech firms had no reason to be confident that they would grow and become able to draw revenues from market sales. Yet their promoters were absolutely convinced that they would succeed and they invested all their resources in their enterprises. Psychologists call it *overconfidence*. More colourfully, Keynes spoke of “animal spirits”. It is a fundamental determinant of entrepreneurship and, possibly, of a lot of decision-making.

Unfortunately, as the example of biotech firms highlights, this kind of conviction may not be based on empirical evidence. Thus, it cannot be captured by evidence theory.

There exists an easy means to overcome this limitation, that consists of defining utility functions on frames of discernment and representing decision-making in very much the same way as expected utility maximisation does. This has been done (Jaffray 1989), but it is not faithful to the rationale of evidence theory.

In fact, Shafer explicitly rejected the idea of separating values from beliefs, which is the basic hypothesis underlying the very concept of utility (Shafer 1986a). To be concrete, Shafer agreed with that ancient Greek tale of a fox who, being unable of reaching grapes,

said that they were sour. Evidently, the utility of the grapes can only be defined if it is independent of the fox's ability to reach them.

Thus, evidence theory calls for tracing the formation of beliefs along the causations of cognitive maps in order to reconstruct decision-making. Shackle had his own suggestion on this topic, namely, that decision-making arises out of focalising attention on two distinct alternatives, to which the final choice is restricted (Shackle 1961). This does not amount to claim that any decision can be thought as an interpolation between two extreme possibilities (Arrow and Hurwicz 1972), but rather points to combining research in mathematical decision theory with experimental psychology.

## References

- Arrow K.J. and Hurwicz L. (1972): “*An Optimality Criterion for Decision-Making under Ignorance*”, in Carter C.F. and Ford J.L. (eds.): *Uncertainty and Expectations in Economics*, Augustus M. Kelley Publishers, Clifton.
- Axelrod R. (1976a): “*The Analysis of Cognitive Maps*”, in Axelrod R. (ed.): *Structure of Decision*, Princeton University Press, Princeton.
- Axelrod R. (1976b): “*The Cognitive Mapping Approach to Decision Making*”, in Axelrod R. (ed.): *Structure of Decision*, Princeton University Press, Princeton.
- Choquet G. (1953-54): “*Theory of Capacities*”, *Annales de l’Institut Fourier*, 5, pp. 131-295.
- De Finetti B. (1931): “*Sul Significato Soggettivo della Probabilità*”, *Fundamenta Mathematicae*, 17, pp. 298-329.
- Dempster A.P. (1966): “*New Methods for Reasoning towards Posterior Distributions based on Sample Data*”, *The Annals of Mathematical Statistics*, 37, pp. 355-374.
- Dempster A.P. (1967): “*Upper and Lower Probabilities Induced by a Multivalued Mapping*”, *The Annals of Mathematical Statistics*, 38, pp. 325-338.
- Dempster A.P. (1968): “*A Generalization of Bayesian Inference*”, *Journal of the Royal Statistical Society B*, 30, pp. 205-47.
- Dubois D. and Prade H. (1988): *Possibility Theory*, Plenum Press, New York.
- Ellsberg D. (1961): “*Risk, Ambiguity, and the Savage Axioms*”, *The Quarterly Journal of Economics*, 75, pp. 643-669.
- Ford J.L. (1987): *Economic Choice under Uncertainty*, Edward Elgar, Aldershot.
- Ford J.L. and Ghose S. (1994a): “*Belief, Credibility, Disbelief, Possibility, Potential Surprise and Probability as Measures of Uncertainty: Some Experimental Evidence*”, University of Birmingham, Department of Economics, Discussion Paper No. 94-22.
- Ford J.L. and Ghose S. (1994b): “*Shackle’s Axioms for Potential Surprise and Belief: A Preliminary Experimental Assessment*”, University of Birmingham, Department of Economics, Discussion Paper No. 94-21.
- Ford J.L. and Ghose S. (1995a): “*Ellsberg’s Urns, Ambiguity and Measures of Uncertainty: Some Experimental Evidence*”, University of Birmingham, Department of Economics, Discussion Paper No. 95-12.
- Ford J.L. and Ghose S. (1995b): “*Shackle’s Theory of Decision-Making Under Uncertainty: The Findings of a Laboratory Experiments*”, University of Birmingham, Department of Economics, Discussion Paper No. 95-08.
- Ford J.L. and Ghose S. (1995c): “*The Primitive Uncertainty Construct: Possibility, Potential Surprise, Probability and Belief: Some Experimental Evidence*”, University of Birmingham, Department of Economics, Discussion Paper No. 95-07.
- Gilboa I. (1987): “*Expected Utility with Purely Subjective Non-Additive Probabilities*”, *Journal of Mathematical Economics*, 16, pp. 65-88.
- Gilboa I. and Schmeidler D. (1993): “*Updating Ambiguous Beliefs*”, *Journal of Economic Theory*, 59, pp. 33-49.
- Good I.J. (1952): “*Rational Decisions*”, *Journal of the Royal Statistical Society B*, 14, pp. 107-114.
- Good I.J. (1962): “*Subjective Probability as the Measure of a Non-Measurable Set*”, in Nagel E., Suppes P. and Tarski A. (eds.): *Logic, Methodology and Philosophy of Science*, Stanford University Press, Stanford.
- Harmanec D. and Klir G.J. (1994): “*On Modal Logic Interpretation of Dempster-Shafer Theory of Evidence*”, *International Journal of Intelligent Systems*, 9, pp. 941-951.

- Hey J.D. (1985): "*The Possibility of Possibility*", Journal of Economic Studies, 12, pp. 70-88.
- Jaffray J.Y. (1989): "*Linear Utility Theory for Belief Functions*", Operations Research Letters, 8, pp. 107-112.
- Jaffray J.Y. (1992): "*Bayesian Updating and Belief Functions*", IEEE Transactions on Systems, Man, and Cybernetics, 22, pp. 1144-1152.
- James G.E. (1996): "*Strategic alliances: Elixers, poison pills or virtual integration? A longitudinal exploration of industry-level learning in biotech*". Unpublished doctoral dissertation, University of Colorado, Boulder.
- James G.E. (2000): "*Industry-Level Learning and the Social Construction of Recipes*", in Sigismund Huff A. and Huff J.O. with Barr P. (eds.): When Firms Change Direction, Oxford University Press, Oxford.
- Katzner D.W. (1986): "*Potential Surprise, Potential Confirmation, and Probability*", Journal of Post-Keynesian Economics, 9, pp. 58-78.
- Klir G.J. and Harmanec D. (1994): "*On Modal Logic Interpretation of Possibility Theory*", International Journal of Uncertainty, Fuzzyness and Knowledge-Based Systems, 2, pp. 237-245.
- Klir G.J. and Yuan B. (1995): *Fuzzy Sets and Fuzzy Logic: Theory and Applications*: Prentice Hall, Upper Saddle River.
- Koopman B.O. (1940a): "*The Axioms and Algebra of Intuitive Probability*", Annals of Mathematics, 41, pp. 269-292.
- Koopman B.O. (1940b): "*The Bases of Probability*", Bulletin of the American Mathematical Society, 46, pp. 763-774.
- Koopman B.O. (1941): "*Intuitive Probabilities and Sequences*", Annals of Mathematics, 42, pp. 169-187.
- Ramsey F.P. (1931): "*Truth and Probability*", in Braithwaite R.B. (ed.): The Foundations of Mathematics and Other Logical Essays, Routledge & Kegan Paul, London.
- Resconi G., Klir G.J., Harmanec D. and St.Clair U. (1996): "*Interpretations of Various Uncertainty Theories Using Models of Modal Logic: A Summary*", Fuzzy Sets and Systems, 80, pp. 7-14.
- Resconi G., Klir G.J. and St.Clair U. (1992): "*Hierarchical Uncertainty Metatheory Based Upon Modal Logic*", International Journal of General Systems, 21, pp. 23-50.
- Resconi G., Klir G.J., St.Clair U. and Harmanec D. (1993): "*On the Integration of Uncertainty Theories*", International Journal of Uncertainty, Fuzzyness and Knowledge-Based Systems, 1, pp. 1-18.
- Sarin R. and Wakker P. (1992): "*A Simple Axiomatization of Nonadditive Expected Utility*", Econometrica, 60, pp. 1255-1272.
- Sarin R. and Wakker P. (1994): "*Gains and Losses in Nonadditive Expected Utility*", in Munier B. and Machina M.J. (eds.): Models and Experiments in Risk and Uncertainty, Kluwer Academic Publishers, Dordrecht.
- Savage L.J. (1954): *The Foundations of Statistics*, Wiley, New York.
- Schmeidler D. (1989): "*Subjective Probability and Expected Utility without Additivity*", Econometrica, 57, pp. 571-587.
- Shackle G.L.S. (1961): *Decision, Order and Time in Human Affairs*, Cambridge University Press, Cambridge.
- Shafer G. (1976): *A Mathematical Theory of Evidence*, Princeton University Press, Princeton.
- Shafer G. (1979): "*Allocations of Probability*", The Annals of Probability, 7, pp. 827-839.
- Shafer G. (1981): "*Constructive Probability*", Synthese, 48, pp. 1-60.

- Shafer G. (1982): "*Belief Functions and Parametric Models*", Journal of the Royal Statistical Society B, 44, pp. 322-352.
- Shafer G. (1986a): "*Savage Revisited*", Statistical Science, 1, pp. 463-501.
- Shafer G. (1986b): "*The Combination of Evidence*", International Journal of Intelligent Systems, 1, pp. 155-179.
- Shafer G. (1990): "*Perspectives on the Theory and Practice of Belief Functions*", International Journal of Approximate Reasoning, 4, pp. 323-362.
- Shafer G. (1992): "*Rejoinders to Comments on 'Perspectives on the Theory and Practice of Belief Functions'*", International Journal of Approximate Reasoning, 6, pp. 445-480.
- Shafer G. and Tversky A. (1985): "*Languages and Designs for Probability Judgment*", Cognitive Science, 9, pp. 309-339.
- Sigismund Huff A. (1990): "*Mapping Strategic Thought*", in Sigismund Huff A. (ed.): Mapping Strategic Thought, John Wiley & Sons, Chirchester.
- Sigismund Huff A., Narapareddy V. and Fletcher K. (1990): "*Coding the Causal Associations of Concepts*", in Sigismund Huff A. (ed.): Mapping Strategic Thought, John Wiley & Sons, Chirchester.
- Simon H.A. (1982): *Models of Bounded Rationality*, The MIT Press, Cambridge.
- Smith C.A.B. (1961): "*Consistency in Statistical Inference and Decision*", Journal of the Royal Statistical Society B, 23, pp. 1-25.
- Smith C.A.B. (1965): "*Personal Probability and Statistical Analysis*", Journal of the Royal Statistical Society A, 128, pp. 469-499.
- Weick K. E. (1990): "*Introduction: Cartographic Myths in Organizations*", in Sigismund Huff A. (ed.): Mapping Strategic Thought, John Wiley & Sons, Chirchester.
- Williams P.M. (1976): "*Indeterminate Probabilities*", in Przełęczki M., Szaniawski K. and Wójcicki R. (eds.): Formal Methods in the Methodology of Empirical Sciences: D. Reidel Publishing Company, Dordrecht.