6th International Symposium on Imprecise Probability: Theories and Applications, Durham, United Kingdom, 2009

How can we get new knowledge?

Frank Hampel
ETH Zurich
Zurich, Switzerland
hampel@stat.math.ethz.ch

Abstract

The paper discusses the (common, important, and yet neglected) situation of a (strong or full) conflict of evidence in scientific and everyday inference (which may lead to valuable new knowledge and even an unexpected scientific breakthrough). It analyses the structure and role of the background knowledge we are using and may have to change, and the many aspects of new information and its interpretation. A number of real life examples follows, which also bring up some more subtle points of inductive thinking.

Keywords. Background knowledge, new information, conflict of evidence, change of model/paradigm, common sense thinking, scientific breakthrough, philosophical foundations of inductive inference, real life examples.

1 Introduction

In observing the various theories using something like upper and lower probabilities, such as in Shafer (1976), Dubois and Prade (1988) and Zadeh (1965), I am still wondering what the precise numerical interpretation of the numbers occurring there is supposed to be, apart from situations with symmetry (cf., e.g., Coolen, 1998) or the start with "total ignorance" which all these theories can deal with (contrary to the neo-Bayesian theory). However, it may well be that only the vague, approximate interpretation of the numerical fixation is relevant, that (like in the Neyman-Pearson theory) in a single situation only values "near" 0 or 1 have a direct practical interpretation, and that perhaps in similar situations the different theories, as far as they are "objective", may lead to somewhat similar values.

A (rare?) example where such a comparison is possible, are enforced fair bets on k independent tosses of a biased coin, starting with total ignorance about the probability of success, evaluated by Smets's pignistic transformation of the Dempster-Shafer belief function theory (Smets, 1990, Smets, 1991, Smets, 1993) and by my own frequentist theory (Hampel, 1993a, Hampel, 1993b, Hampel, 1998, Hampel, 2001; cf. also Hampel, 2002, Hampel, 2005). The enforced probabilities of $(0, \ldots, k)$ successes are for $k = 1$ (1/2, 1/2) (for both and many other theories), for $k = 2$ (5/12, 2/12, 5/12) (Smets) and (1/2, 0, 1/2) (Hampel), for $k = 3$ (157/432 = 0.363, 59/432=0.137, 59/432, 157/432) (Smets) and, for the symmetric solution, (5/12 = 0.417, 1/12 = 0.083, 1/12, 5/12) (Hampel). The numbers are clearly different, but still show some superficial similarity.

Exact numbers may be needed in intermediate calculations (to avoid rounding errors), and they are important in well-developed quantitative theories, where the aims are "only" numerical refinements within a given frame. (The "only" should not be misleading; most research is of this type, and also within a fixed frame there is qualitative progress possible, as by tests.) But when I look at everyday learning and also at scientific breakthroughs, I find that progress often comes by abandoning an old framework or paradigm and replacing it by a new one (cf. also Kuhn, 1962). Such a replacement should obviously be considered when there is a contradiction between the old framework and a new observation.

But in the literature (as far as I know it) I find discussion of this model change conspicuously absent. Only top applied statisticians like John Tukey or Cuthbert Daniel dare to "change the horses in the middle of the stream" (C.D.). The Neyman-Pearson theory is very anxious not to change the assumed model, because then some probabilities would be changed; but these probabilities may have become completely irrelevant. Neo-Bayesians renormalize their posteriors, no matter how small they are without renormalization. For example, depending on circumstances, a single outlier can play havoc with their results. (The (in)famous

dictum by de Finetti: "There is no Bayesian problem of outliers, because there are no outliers" assumes an omniscient, God-like attitude: in real life, we just don't know everything.) Shafer (1976) discusses at length a quantitative measure of the degree of contradiction between two claims; but he does not say what to do when the contradiction is large (except renormalizing) or even complete.

Now one reason for the silence about such a crucial point is probably that it may be (or may seem) impossible to build an exact, numerical mathematical theory about it. However, my impression is that most arguments in real life and many arguments in top science are even on a 0-1-scale (not involving degrees in between), but involving changes of background beliefs; therefore I consider it worthwhile to study the (existing) qualitative (and perhaps even semi-quantitative) structures which we can find when we analyse our corresponding thinking in more detail. Such an attempt is what the paper is about.

This paper is closely related to the short outline in Hampel (2007); cf. also Hampel (2009). After an introductory example, the interplay between background knowledge and new information, the structure of background knowledge in real life, and the interpretation and surrounding structure of new information are discussed; then a number of real life examples are given, including the discussion of some finer points in inductive logic.

While in Shafer (1976) the two (or more) sources of new information are treated symmetrically, keeping the background model fixed, here one of the sources of information is the background model itself which may have to be changed by the new information.

Although not discussed here, the new concepts can easily be applied to the problem of model change in applied statistics or data analysis, using the great experience of top applied statisticians.

We also note that probably quite a few scientific breakthroughs, like the discovery of penicillin by Fleming, or the discovery of the effect of rubella on pregnancy, have their root in an unexpected (and first unexplainable) observation which was then thoroughly analyzed.

One referee kindly alerted me to the danger that my approach might be confused with the (relatively popular) work on "belief revision" as exemplified in the classic paper by Alchourron et al. (1985). But that paper deals merely with the logical consequences if in a complex logical system one statement is being contradicted (or another statement is being added). This is certainly a legitimate topic of research, but it is entirely restricted to deductive logic, working out the (intricate) logical consequences of partial knowledge. By contrast, I am considering the situation that a former belief is *entirely* wrong, and a new belief has to be created on the basis of inductive guesswork (based on "life experience"). It is one of my main points that such arguments cannot be derived by pure deductive logic (except, of course, if one believes to be omniscient, like God or some Bayesians). Nevertheless, I do describe a rich new structure of inductive thinking. And while my paper abounds in real life examples, I cannot find a single real life example in the 20 or so pages of Alchourron et al. (1985). (In addition, I believe that in practice often the contradictory new observation in their paper needs one or more detailed interpretations in order to allow meaningful logical deductions.) Overall, I think that my approach is often much closer to real life problems (and to Kuhn, 1962) than the approach in Alchourron et al. (1985).

2 Oregon and Dolomites

The following story (which may well be more widely known) was told to me in 1984 by the late Philippe Smets.

A couple (perhaps from the US East Coast) was planning their holiday travel. The wife had found a very enticing article about the Dolomites in a travel journal and wanted to go there. But the husband found in the same journal a highly commendatory article about a dry and sunny place in Oregon, and everybody knows that in Oregon it always rains. Later, the husband found out that there are indeed dry and sunny places in Oregon, and both went to the Dolomites.

Let us now analyze this little story in more detail.

The wife got and accepted a "new information" from the travel journal, namely that the Dolomites would be a very nice place to visit. The husband was more sceptical, but since he could not directly judge the article on the Dolomites, he tried to get some more general "information" on the overall reliability and quality of the travel journal, from which to "extrapolate" to the Dolomites article. And he found an article (on some place in Oregon) which was in contradiction with his "background knowledge." Believing his "background knowledge" more than the unknown quality of the journal, he at least cast some doubt on the praise of the Dolomites. (Or perhaps he found the Oregon contradiction just by accident; the end result would be pretty much the same.)

But then he happened to learn (reliably) about the cold desert in the thinly populated (and hence often forgotten) East of Oregon, refining and in this par-

ticular case correcting his "background knowledge"; thus there was no contradiction and no reason to mistrust the journal anymore, and the couple decided to trust the recommendation for the Dolomites. (It could even be argued that a journal talking about dry spots in Oregon is rather sophisticated and not just citing mainstream beliefs and hence trustworthy, once the existence of these dry spots is acknowledged.)

3 Background knowledge and new information

I think this story is an (already somewhat intricate) example of the following general scheme. We all have accumulated, throughout our lifetime up to the present, a large body of "background knowledge" which we use, often subconsciously, to judge our present surroundings. (The structure of our background knowledge is itself very interesting and important, see Section 4 below.) When we now get some "new information" (be it by words, by experiment, or by observation and experience), we compare it with the "extrapolation" from a pertaining part of our background knowledge; if there is a contradiction, then (apart from the chance of later getting new, clarifying information) we have to dismiss either the new information or some part of our background knowledge; or at least we have to "reinterprete" the one or the other or both, in order to make them compatible again.

This is the classical, rational, "scientific" procedure. However, we might also try to live with a contradiction in our "new background knowledge"; and this not only due to irrational or confused thinking, or in fields like religion ("credo quia absurdum"), but also in pure science such as quantum mechanics (like in the saying about the physicist who believes light is a particle on Mondays, Wednesdays, Fridays, and a wave on Tuesdays, Thursdays, Saturdays, and on Sundays he prays).

A rational way of living with a contradiction is to transcend two contradictory claims A and B by not believing either A or B, but by merely noting that both claims exist, without committing oneself. This is possible in pure inference, as opposed to decisions, in view of the necessary action there; but as long as no action is necessary, and even after a necessary action, it makes sense to consider both A and B possible (as can be done in all theories with something like upper and lower probabilities, beliefs, and so on). If later we are reliably told that the chances, or something similar, of A over B are 999:1, then we most strongly keep A, even if before we had made the (apparently bad, on hindsight) decision B. A theory (like the Bayesian one) which would make any decision, however shaky, automatically part of our new background knowledge, would (with a suitable formalism) still keep B; it would weigh internal consistency over time (!) higher than eventual truthfulness.

4 Background knowledge and real life

But where does our background knowledge come from? It has a wide variety of sources: wishful thinking; prejudice; emotions; belief from hearsay, especially from "authorities"; belief from the media, including the internet; a more or less detailed "official" scientific knowledge (which, as experience shows, is mostly fairly stable, but still continuously and sometimes even fundamentally revised, and which contains many bold extrapolations which are hard to judge for the outsider and which may well turn out to be false); personal experiences and extrapolations from these, in combination with the "official" knowledge, and a broad mix of all these sources. Many fundamental beliefs and attitudes go back to our education and even heredity; but we leave this to others to discuss. However the clearer we know the sources of our convictions, the better we can deal with conflicting new information, in judging the relative reliability of both claims.

The background information comes in layers. Usually we take only the most obvious layer or belief for granted, and when the new information is in agreement with it, then this belief will only be somewhat reinforced. But when there is a contradiction between default background and new information, then we have to dig deeper and choose a less likely background as our updated background. Even then, we shall usually consider only interpretations of the next most likely layer (there may be several).

The idea of looking at the set of all possible interpretations of the world and then choosing the most likely one may sound good in philosophy and in pure mathematics, but this is not how we work in real life, nor in science. A physicist will not consider the set of all possible physical theories and then select the most plausible one, once he is forced by experiment to abandon an old theory; but rather he will only look at a "neighborhood" of the old theory and try to get along with as few (or as simple) changes as possible (simple changes may be radical, but only as much as needed by circumstances). If we wanted to consider ALL possibilities of what could happen when we leave the house (like the famous tile dropping from the roof; or being shot to death by mistake, as happened to a wellknown statistician in Mexico City), we would never set a foot in front of the door. This is

a question of efficiency of life. We normally act and think as if only the most likely, or "most plausible" assumptions would be true.

In addition, we may also look at the set of alternatives which are still "quite possible" (like an unexpected delay in something), just to be on the safe side, depending on how pessimistic we are or how strong the consequences would be. But if we observe a contradiction with the "most plausible" assumption, we fully switch to the set of "quite possible" alternatives and perhaps choose the most likely one among them. Only when an occurrence would have drastic consequences (as in cases of life and death), shall we look also at "unlikely" events (and perhaps write a testament or take out an insurance). We hardly ever (except in theory) shall consider "extremely unlikely" interpretations of the world (or even, for logicians and pure mathematicians, "impossible" ones).

These ordered categories: "most plausible", "quite possible", "unlikely", "extremely unlikely" (and perhaps "impossible"), of which we normally only use the first and the second one, to me seem to provide a sufficiently accurate, but also important and necessary valuation of aspects of reality, both in science and in real life. These valuations may differ according to personal experience and present circumstances (cf. the examples below). New experiences may change the category, but usually only to a neighboring one. Cf. also Hampel, 2007.

5 New information

The "new information" in general is not simple and unstructured either. It is connected with "everything that can be said about it", by considering it from its meaning, its sources, its context, its aims, its different possibilities of interpretation, and so on. When faced with new information (and the problem of reconciling it with the background information), and if "So what?" is not the most appropriate reaction (it often is!), the following questions may be helpful:

Who says so?

What is the purpose behind it?

What says the other side? (If controversial)

How does one know this?

What is lacking? (What was forgotten or concealed?)

What does this really mean?

The reliability of the source of information is clearly very important. (I once studied and compared two locally wellknown newspapers for a while. One had a surprisingly large number of – mostly small – inaccuracies. The other, supposed to be very reliable, was so most of the time, but sometimes it contained big blunders – the more misleading as they were unexpected.)

The purpose of news may be a "good story", the fame of a scientist (and the associated money), the need to publish something rather than perish, political influencing, but also neutral information, like the weather report. (Even the weather, and more so the climate, can be political, and even in leading Western countries sometimes scientists have been forbidden to publish their findings.)

The old Roman rule: "Audiatur et altera pars! Listen also to the other side!" is very important in all controversial issues. A comparison of the arguments, motives, backgrounds, reputations, etc. may well allow a decision for one side or the other. Often the truth is somewhere in between; sometimes it is even beyond the range of present opinions.

The question how the new information could have been obtained means going beyond the surface of the information to its possible origins. Sometimes these origins are very subjective and biased, or shaky in other ways.

As is wellknown among statisticians (and still not enough known among nonstatisticians), every statistical number should have with it at least an implicit rough indication of its statistical accuracy. But this is not enough. A good, objective information should also contain a discussion of possible systematic (and semisystematic) errors and their orders of magnitude, of likeliness and effects of gross errors, and of possible reinterpretations of the findings, which might show the results in a completely different light. And often we can only hope that no relevant information has been left out of the discussion. We are reminded of the (in)famous "oath of the statistician": "I swear to tell the truth – nothing but the truth – but not the whole truth." Contrary to deductive logic, conclusions in inductive logic can be changed completely by leaving skillfully out part of the premises.

Often it pays to go a step back and ask oneself: Is this information really what it is supposed to be? Or does it actually mean something noticeably different? Is it only suggestive, and perhaps even without real contents?

A delightful collection of arguments and examples in these directions can be found in the classic book "How to lie with statistics" (Huff 1954); there is also a number of more recent books along similar lines.

6 Some examples of interpretation of new information

A prototype situation is the following: We are living on, without much thinking, in our "most plausible" world, and then (if we are awake and attentive) we observe something strange and surprising (like a Zurich tram in the wrong street, cf. Hampel 2007), which forces us to consider other interpretations of our present surrounding reality (e.g., an accident) which may influence our plans (e.g., requiring us to take another route, or enforcing a delay). Thus, several "quite possible" interpretations are raised to the category "most plausible", until we have learned more. The event observed might even be a "non-event", like suddenly no cars coming from the opposite direction, or the not-barking of the dog in a story about Sherlock Holmes.

The following examples are often from my own experience, especially from ornithology: because I know them best, and because they are sufficiently "unimportant" to allow a neutral discussion. If we discussed "God and the World", which we formally could do equally well, we would soon end up in heated arguments about "God and the World" and not the logical structure of thinking.

6.1 Alarms

When I recently heard a siren wailing at home, I remembered that there were regular test alerts, but I did not remember when. So I looked at the watch; the "round" time (precisely a half-hour) seemed to confirm this interpretation, but to be more sure, I looked also out of the window to see the people on the street walking casually as usual. (The newspaper announcement of the trial alert, which I found later, was somewhat hidden.)

A real alarm under my circumstances fortunately was not very probable, but one never knows. For me, in the beginning both real and test alert were "most plausible" (though the test alert was much more "probable" in the subjectivistic Bayesian sense), and only the two indications (and later the proof) diminished the plausibility of a real alarm.

However, the year before, there was a real alarm in a nearby community because of a pollution of the drinking water. Since it was not too long after the test alert, many people did not pay any attention to it. (In addition, the alarm came only more than five hours after the pollution; and many people were sick for several days.)

Some years ago, during the wars in ex-Yougoslavia, a child from that region had come to Switzerland and went to a Swiss school. When an airplane flew low over the school building, this child immediately dove under a table; and the Swiss classmates had to learn how lucky they were not to be traumatized in this way and not carrying such a background experience with them.

6.2 The meaning of a phrase

The interpretation of a new information may depend strongly on the context. Thus (cf. Hampel 2007), when we ask someone whether a certain way leads to a certain place, in our Western culture a "yes" normally just means "yes" (unless there are or may be reasons that the person answering may want to lead us astray). But experienced world travellers have gained the background knowledge that a "yes" in a different cultural context can mean many different things, for example: 1. "Yes". 2. "Yes, I understand your question." (Perhaps the actual answer comes later.) 3. "Yes, I heard that you said something (without understanding it)." 4. "Yes – you seem to believe so, and I don't want to contradict you." 5. "Yes – I really don't know." 6. "Yes – any other answer would be impolite." (Cf. the story of the East Asian student in Berkeley who finally learnt to say "yes yes" or "yes no", depending on what he really meant, because he was obliged to say always "yes.")

(There is also the true story of the white man who spoke perfectly well Chinese and who asked two old Chinese men whether this was the way to the Ming graves. The two men just stared at him openmouthed; he asked again; the same reaction. Finally he gave up. When he was leaving, he heard one man say to the other: "This sounded just as if he asked whether this was the way to the Ming graves." These men certainly had a strong background conviction.)

More generally, let us assume we learn that a person makes a statement "A". This may mean: 1. A; 2. the opposite of A; 3. approximately A; 4. perhaps A; 5. something related to A; 6. A and B (A incomplete and misleading without B); 7. a polite phrase with no other meaning; 8. an attempt to conceal B, and to divert attention away from B (a frequent trick of tourist guides, if B would be embarassing); 9. an unsubstantiated claim (advertisement); 10. a misunderstanding or a mistake: what was meant was B; 11. A under a side condition forgotten to be mentioned; 12. A under an assumption obvious to the speaker, but unfortunately wrong; 13. A and a seemingly obvious conclusion B, which however is wrong; 14. A and the denial of a conclusion B which is considered too obvious to be true (as pure mathematicians sometimes think); 15. A, but not a fully obvious and

correct conclusion B (which would cause a judge to be called prejudiced and biased; this problem seems to be not uncommon in law), and so on.

(We are also reminded of the joke about the absent-minded professor who says A, writes B, thinks C, means D, and E would have been right.)

6.3 Prejudices I and overreactions

There is often a tendency to cling to old convictions and to defend them by exaggerated means. When I once in fall discovered a Citril Finch (Serinus citrinella) in the Harz mountains in northern Germany, far north of the nearest breeding range in the Black Forest which it hardly ever left, suddenly the Citril Finch was supposed to be a "rather common cage bird" (which it definitely was not, though it was entirely appropriate to consider the possibility of an escaped bird). But when some weeks later I discovered a whole flock of Citril Finches in the same area, opinions switched to the other extreme that some ornithologists believed the bird was even breeding in the Harz. (Compare also the extreme switch of opinions about redescending M-estimators in the Princeton Monte Carlo study, cf. Hampel 1997.)

A rather ridiculous attempt to defend a preconceived attitude by all means once happened in Zurich, when many people saw an "UFO" (a slowly descending chain of lights) in about 10 km distance in a very hazy night. Since some explanation had to be found (to dispel any chance of believing in the little green men), this was officially explained (and believed, even by hobby astronomers) as a chain of burning candles hanging below balloons! As I explained in my farewell lecture, it was nothing but a chain of car headlights descending from an (invisible) mountain lookout.

A very illuminating experiment concerning the strength of false imagination, but also the occurrence of rare exceptions, was once done by an astronomer on British TV (Hunt & Moore 1982, p. 32f). Near inferior conjunction of Venus, he showed the telescopic view of its crescent, whose visibility with the naked eye (under favorable circumstances) is a question of debate among astronomers, and asked the viewers to send in little sketches when they thought they saw the crescent. More than 200 sketches were received; all but two – both by surprised young people – showed the crescent in the inverted view of the telescope. Apparently only these two people genuinely saw the crescent of Venus (as is corroborated by a number of other well-documented observations). Thus, 99% of the claims were illusions, but 1% were proper.

6.4 Layers of questioning

A sceptic (who does not know about the other evidence) might still claim that the two young people could be cheating: they could have known about the inversion of astronomical telescopes and, to make it look more convincing, they might have claimed to be surprised about the right picture (which they did not see). In our case, this appears to be a very far-fetched argumentation, especially since the stake is very low; but in other situations, sceptical digging into deeper layers might well be appropriate.

During the period when I was collecting the bird observations in southernmost Lower Saxony, a young field ornithologist claimed to have seen a female Red-crested Pochard (Netta rufina) on a certain lake, which would have been only the second record of this (mostly very rare) duck for the whole area. I let him describe his observation in a neutral mood and asked him also whether he could see the little red spot at the bill of the female. "Oh yes", he said, "the sun was so bright that it looked as if the whole bill was red." Then I knew two things for certain: that it was a male Red-crested Pochard in eclipse plumage (which, as I knew, has an all-red bill and otherwise looks like a female, but which was not painted even in the best bird book of that time), and (as I had not doubted anyway) that the observation was not made up. (It was also to the credit of the observer that he was very aware of the dangers of light effects.) It can often pay to have some more knowledge or experience than the other person. And what was puzzling for him (the bright red bill), found an explanation and was a proof that the observation was basically correct (apart from the sex).

When I prepared a talk for the European Meeting of Statisticians 1987 in Thessaloniki, I also read something from Aristotle, the genius loci, and was surprised to find that what we consider his "logic" was only a small part of his discussions of a logic in a much broader sense. One of his examples was the story of a very strong man who was accused to have robbed another person during the night. His defense was that he would never have done it, because if he did, he knew that he immediately would be suspected and arrested (being the strongest man around). So it could not have been him.

We can iterate this argument: Since he had such a convincing (?) defence, he could have been the robber, after all. (Cf. the "Theorem" 2 in my talk in Thessaloniki: "The game is indefinite.") Where to put the limit and stop? In general, this may be a difficult problem, with no guidelines except insight into the situation and common sense. In the cases of the cres-

cent of Venus and of the Red-crested Pochard, there clearly was no reason to go further (also because I knew the observer personally in the second case), but when the stakes are high, the question becomes more delicate. The stake might, for example, be the fame of some sort, as in the case of the British ornithologist who shot birds in Asia, imported them frozen to England and layed them out in a small stretch of sea shore where he then obtained "first records" and other remarkable "rare records" of these species for Great Britain (even with "proofs", namely the dead bodies).

He was actually convicted by a statistical argument: there were far too many "rare birds" concentrated on that otherwise rather ordinary piece of sea shore, also compared with the wider surroundings. But even here one has to be careful. While on suitably located islands like Heligoland or the Scilly Islands many rare bird records can be expected, the number of special records around Hildesheim (Lower Saxony), in a very "ordinary" landscape, is at first really amazing. In this case it was due to the sheer fanatic ardor with which the Hildesheim group of ornithologists made their (well-documented) observations; and it showed how little we really know about our surroundings.

6.5 Subtle clues

It is a general experience of mine that unintentional, casual, neutral, often subtle observations or remarks often have the "ring of truth" (as long as this is not used against me on purpose, in the next round of argumentation as described above), while I mistrust all claims with a hidden (or even obvious) purpose behind them. An example of my beginner's time in ornithology is a flock of (very variable) Dunlins (Calidris alpina) in fall; as I counted them back and forth, each time my eye stayed longer with a particularly clean bird (which in a process of "Gestaltwahrnehmung" seemed more and more like a nearby outlier), until I flushed it and could safely identify it as a Curlew Sandpiper (Calidris ferruginea), an uncommon migrant from Siberia.

It may also be that something seems "to be the same and not the same", as when I twice in 3 days observed a Kentish Plover (Charadrius alexandrinus), which is very rare inland. At closer scrutiny one was a male and the other a (distinguishable) female.

Sometimes also "traces of memory" can be helpful for explaining a strange observation.

A very informative clue can be the "Gestalt" of a bird song. Once I woke up by a bird song I had never heard before; it was a Greenish Warbler (Phylloscopus trochiloides), one of the first records in Western Germany, later published (Hampel 1964; Hampel 1965) and corroborated by several other West German records during the same summer. Decades later I was thrilled to hear and recognize the same song again for the second time in the wintering area in India.

Another acoustic observation was more complicated. On May 31, 1985, just before leaving Poland, I heard a new song at Milicz railway station (Silesia) which according to the Swedish bird records appeared to be an Arctic Warbler (Phylloscopus borealis). But the scientists I contacted claimed that the Wood Warbler (Ph. sibilatrix) can have a very similar song. So I spent some summers to check the breadth of variation of Ph. sibilatrix songs. There was some variation, but I never came close to the song in question. (Of course, I cannot exclude that in some areas Ph. sibilatrix can sing almost like Ph. borealis, but I suspect that it was the same reaction as with Serinus citrinella suddenly being a "rather common cage bird.") Meanwhile, I got a record with Mongolian bird songs, including several songs by Ph. borealis, and one sounded exactly like the bird I had heard. My last personal doubts vanished when in the tropical jungle of southwestern China, amidst lots of new songs, I suddenly heard again the Milicz song (and briefly saw the bird).

I got some feedback on my observation of Oct. 20, 1962, of a possibly Phylloscopus schwarzi in Goettingen (Hampel 2007), asking why I did not put it from the category "extremely unlikely" to "possible" if not "plausible", but only to "unlikely" after hearing of the "invasion" in Europe. But in this case I had very little positive evidence for the species. I mainly knew that according to the call it was not Ph. collybita (nor Ph. trochilus), but I could not positively and safely identify the call, not knowing more about this and other similar-looking Asiatic accidentals (and about details of the "invasion"). Nevertheless, with additional information a new assessment might be possible.

6.6 Prejudices II and stability of opinion

It is a hard situation when people are so convinced of their "background knowledge" that they refuse to look at anything nonfitting (like the astronomer and the philosopher in Brecht's "Leben des Galilei" who refused to look through Galilei's telescope with the moons of Jupiter visible, only arguing whether such moons were "possible" and were "necessary"). I once had an experienced ornithologist with me who literally (for a long time) refused to look at a Crane (Grus grus) who was there at a very unusual time of year, because he "knew" it could not be. (Probably he thought I was pulling his leg.)

There is actually a fairly unknown variant of Bayes'

theorem, derived from general principles, with an exponent on one of the two factors. As the exponent varies between zero and infinity, we get all kinds of people from those who are completely stuck in their prejudices, to those who believe everything. (A fitting story is the Sufi story of two persons A and B arguing strongly; a third person C listens to A and says: "You are right." Then he listens to B and says: "You are right." When another person points out to C that A and B cannot be both right, he says: "You are also right.")

It is clear that some medium stability of opinion is needed in the flow of new informations, and science certainly should lean somewhat more to the conservative side. But when a reputable scientist observed, with good documentation, that Lichtenstein's Sandgrouse (Pterocles lichtensteinii), an extreme desert dweller, flew daily up to 80 km to the nearest waterhole, walked into the water until the belly feathers soaked up the water like a sponge, and then flew back to water the young birds in the nest with its belly, this was first ignored and then emphatically denied for 70 years, until it was confirmed also for several other Pterocles species (cf. Scott et al. 1974, p. 153).

Another such story (cf. Barth 1991): the similarity of orchid flowers of the genus Ophrys with several species of sand bees had long been noticed; but when a wellknown specialist observed an actual "copulation" attempt between bee and flower, he first kept it for himself; and when he later wanted to publish it, it was put down as "dirty fantasies of an old man." (By now, there are not only documentary movies, but also fascinating research about the female smell of unpolluted and polluted flowers, as well as a new systematics of the orchids based on the bees.)

6.7 Some tough situations

Very often we have to deal with half-truths ("there may be something to it ...") which are very hard to judge properly.

But one of the worst things that can happen to the pursuit of truth in science is when it is distorted and suppressed by political and religious, commercial and financial interests, as happened again and again. In evaluating new evidence (or even the lack of public evidence), we unfortunately have to take such interests and influences into account.

Acknowledgments: I owe to Werner Stahel, besides technical help, several valuable remarks. – One anonymous referee provided me with an additional reference.

References

Alchourron, C. F., Gardenfors, P. and Makinson, D. (1985). On the logic of theory change: partial meet contraction and revision functions, *The Journal of Symbolic Logic* **50**(2): 510–530.

Barth, F. G. (1991). *Insects and Flowers: The Biology of a Partnership*, Princeton University Press, Princeton, N.J.

Coolen, F. P. A. (1998). Low structure imprecise predictive inference for Bayes' problem, *Statistics and Probability Letters* **36**: 349–357.

Dubois, D. and Prade, H. (1988). *Theory of Possibility*, Plenum, London, UK. *Original Edition in French (1985) Masson, Paris.*

Hampel, F. (1964). Grüner Laubsänger (Phylloscopus trochiloides) in Göttingen, *Journal für Ornithologie* **105**(2): 199.

Hampel, F. (1965). Artenliste vom Seeburger See 1955-1964 (unter knapper Berücksichtigung des Raumes um Göttingen), Mimeographed manuscript, Göttingen, 23 pp (later reprinted).

Hampel, F. (1993a). Some thoughts about the foundations of statistics, *in* S. Morgenthaler, E. Ronchetti and W. A. Stahel (eds), *New Directions in Statistical Data Analysis and Robustness*, Birkhäuser Verlag, Basel, pp. 125–137.

Hampel, F. (1993b). Predictive inference and decisions: Successful bets and enforced fair bets, *Proc. 49th Session of the ISI, Contrib. Papers, Book 1. Firenze (Italy)*, pp. 541–542.

Hampel, F. (1997). Some additional notes on the "Princeton Robustness Year,", *in* D. R. Brillinger, L. T. Fernholz and S. Morgenthaler (eds), *The Practice of Data Analysis: Essays in Honor of John W. Tukey*, Princeton University Press, Princeton, pp. 133–153.

Hampel, F. (1998). On the foundations of statistics: A frequentist approach, *in* M. S. de Miranda and I. Pereira (eds), *Estatística: a diversidade na unidade*, Edições Salamandra, Lda., Lisboa, Portugal, pp. 77–97.

Hampel, F. (2001). An outline of a unifying statistical theory, *in* G. de Cooman, T. L. Fine and T. Seidenfeld (eds), *Proc. of the 2nd Internat. Symp. on Imprecise Probabilities and their Applications, ISIPTA'01, Cornell University, 26-29 June 2001*, Shaker Publishing Maastricht, 2000, pp. 205–212. *Also:* `ftp://ftp.stat.math.ethz.ch/Research-Reports/95.pdf`

Hampel, F. (2002). Some thoughts about classification, *in* H.-H. Bock, K. Jajuga and A. Sokołowski (eds), *Classification, Clustering, and Data Analysis. Recent Advances and Applications*, Studies in Classification, Data Analysis, and Knowledge Organization, July 16–19, 2002, Cracow, Poland, Invited keynote lecture, 8th Conference of the International Federation of Classification Societies, Springer, Berlin, pp. 5–26. *Also:* `ftp://ftp.stat.math.ethz.ch/Research-Reports/102.pdf`

Hampel, F. (2005). The proper fiducial argument. Extended abstract, *Electronic Notes in Discrete Mathematics* (Elsevier Science) **21**: 297–300.

Hampel, F. (2007). Upper and lower probabilities in real life, *CD-ROM containing the Proc. 56th Session of the ISI, Contrib. Papers, Lisboa, Portugal*. *Also:* `ftp://ftp.stat.math.ethz.ch/Research-Reports/145.pdf`.

Hampel, F. (2009). Nonadditive probabilities in statistics, *Journal of Statistical Theory and Practice* **3** (No.1: Special Issue on Imprecision): 11–23. *Also:* `ftp://ftp.stat.math.ethz.ch/Research-Reports/146.pdf`

Huff, D. (1954). *How to Lie With Statistics*, Lowe and Brydone, London, UK.

Hunt, G. E. and Moore, P. (1982). *The Planet Venus*, Faber and Faber, London.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*, University of Chicago Press.

Scott, P., Fry, C. H., Flegg, J. J. M., Vevers, G. and Pettingill Jr., O. S. (eds) (1974). *The World Atlas of Birds*, Crescent Books, New York.

Shafer, G. (1976). *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, N. J.

Smets, P. (1990). Constructing the pignistic probability function in a context of uncertainty, *in* M. Henrion, R. D. Shachter, L. N. Kanal and J. F. Lemmer (eds), *Uncertainty in Artificial Intelligence*, Vol. 5, Elsevier Sci. Publ., pp. 29–39.

Smets, P. (1991). The transferable belief model and other interpretations of Dempster-Shafer's model, *in* P. P. Bonissone, M. Henrion, L. N. Kanal and J. F. Lemmer (eds), *Uncertainty in Artificial Intelligence*, Vol. 6, Elsevier Science Publ., pp. 375–383.

Smets, P. (1993). No Dutch book can be built against the TBM even though update is not obtained by Bayes rule of conditioning, *in* R. Scozzafava (ed.), *Workshop on Probabilistic Expert Systems*, Soc. Italiana di Statistica, Roma, pp. 181–204.

Zadeh, L. A. (1965). Fuzzy sets, *Inform. Control* **8**: 338–353.