When Laypeople are Right and Experts are Wrong: Lessons from Love Canal

Ragnar Fjelland

Abstract: Love Canal, a suburban town in New York State built on a waste disposal site of a former chemical factory, provoked one of the first major environmental controversies. It involved scientists, citizens and politicians, including the US Congress and President. The controversy raises many important problems, and the article focuses in particular on the uses of scientific knowledge and the role of scientists. Although the scientists worked for the authorities, they regarded their knowledge as objective and their advice as neutral. However, the residents of Love Canal did not trust them, and engaged their own scientist. At the time of the controversy (1978) the Precautionary Principle had not been formulated, but the controversy involved many issues that have later been related to the principle. One particular issue was the uses of statistics, and the relationship between type 1 and type 2 statistical errors. The article relates the controversy to recent debates on the proper use of significance tests and statistics, and argues that context and values have to be taken into consideration. It concludes that in cases like Love Canal it is imperative to inform about uncertainty and to involve all stakeholders.

Keywords: *environmental controversy, precautionary principle, statistics, uncertainty, expert versus non-expert knowledge, ethics of science.*

1. Introduction: Why the Love Canal controversy is important

Love Canal is a suburban town in the US state of New York close to Niagara Falls. In the summer of 1978 it became the scene of one of the first major environmental controversies, a controversy that hit the headlines worldwide. The residents complained that chemicals from a nearby chemical waste site caused health problems. Love Canal was in fact built over the waste disposal site of a former chemical factory. Nearly 22,000 tons of chemical waste (including polychlorinated biphenyls, dioxin, and pesticides) had been deposit-

ed. The controversy involved scientists, experts, and politicians, including the governor of the State of New York, the US Congress and the President.

There are many lessons to be learned from the case. Many of them concern the relationship between the authorities, the affected citizens, and the general public: The importance of giving honest information and not downplaying the hazards, and of recognizing the affected citizens as partners and resources, and not just as ignorant and hysterical people. Other lessons concern the uses of scientific knowledge and the role of scientists. I shall in this article mainly concentrate on lessons we can learn from Love Canal about the role of science and scientists.

I will focus on the crucial question: Was there a causal relationship between the toxic chemicals in the waste site and health problems among the residents of Love Canal? To put it simply: Was Love Canal a safe place to live in? New York State Department of Health engaged scientists to give an answer to that question. The scientists regarded it as their job to find the facts, and leave it to the authorities to make decisions. Although they reported to the authorities, they regarded themselves as objective and neutral. However, the residents in Love Canal did not trust them, and they themselves took an active part and engaged other scientists.

Were the scientists objective? The question does not only concern the objectivity of scientists, but the deeper question of objectivity itself. The traditional view, held by the scientists working for the Department of Health, was that facts should be separated from values, and that scientists should stick to the facts. I shall in this article show that their ideal of objectivity led to bad science, and that it implicitly favored the authorities.

The questions raised in Love Canal are today more urgent than ever, because many of the problems contemporary science faces have similarities to the Love Canal case, characterized by complexity and uncertainty.

2. The story

The story of Love Canal began in the 1890s, when William T. Love planned to build a hydroelectric power plant and started the construction of a canal to supply water from the Niagara River. However after a section of the canal, which was approximately one kilometer long, twenty meters wide, and three meters deep, had been dug, a financial crisis forced the project to be abandoned. In 1905 Hooker Electrochemical Company was started near Niagara Falls. The factory produced, among other things, chlorine and caustic soda. Between 1942 and 1952 it was allowed to dispose of almost 22,000 tons of chemical waste in the canal, in fiber and metal barrels. In 1953 the canal was

full, and was covered with soil and clay, while grass grew on the surface. The school board of Niagara Falls then bought the land, including the canal, for one dollar. One condition of the deal was for a disclaimer to be included in the contract exempting Hooker Company from any future liability (Levine 1982, pp. 11ff.).

The school was completed in 1955, with a capacity of 400 children, who attended every day. Houses were built around the school, most of them modest two- and three-bedroom houses. Although there had been some focus on the chemicals when the school was constructed, most of the residents were unaware of the chemicals that were buried in the ground. And although there were signs of leakage early on, the residents did not pay attention to them. However, things became worse in the 1970s. For example, after heavy rainfall odors from chemical substances became noticeable. The first investigation was carried out in 1976, and the suburb was visited several times by officials of the New York State Department of Health. However, in the summer of 1978, when the newspapers made the general public aware of the situation, the Department of Health was forced to act. In their preliminary investigations they found more than eighty different chemicals at the waste site. Ten of these substances were known to be potentially carcinogenic (*ibid.*, p. 41). The director of the Department of Health declared a state of emergency due to the danger to public health. He justified this on the grounds he was convinced that the toxic chemical substances from the waste site represented a danger to the residents in the area. Soon after that the governor of New York offered to buy the 239 houses situated nearest to the waste site and to help the residents relocate. A fence was put around the evacuated area. The Department of Health began an investigation into the health of the residents, this included blood tests, questionnaires, and an inquiry into the incidences of ill health among the inhabitants. Early in the Fall of 1978 the preliminary results of the study were made public; the health authorities assured the residents that the rest of Love Canal was a safe place to live (Paigen 1982, p. 29).

However, the residents distrusted the information given by the authorities. They felt that the authorities regarded the residents themselves as the problem and suspected that information was being withheld from them. The residents were also never consulted during the process as key contributors of information (Levine 1982, p. 27). One of the residents was Lois Gibbs. She was a young housewife who had neither been interested in politics nor had taken part in any organized political activity. However, she started organizing her neighbors into what became Love Canal Homeowners Association (LCHA), and became its leader. They assisted in the investigations of the Department of Health, among others by calling neighbors and urging them to fill in forms, and by registering medical ailments. One night Gibbs decided

to go through all the material that had been collected. She sat down with a map of Love Canal and put a pin on every house that had registered medical problems. It appeared that the pins formed a pattern of narrow paths on the map. Older residents had previously informed her of large stream-beds and swales cutting through the area. They had been filled in when the houses were built. It turned out that photographs from the 1930s displayed the original swales and stream-beds. Gibbs had the idea that there might be a connection between the pattern of the pins on the map and the swales (*ibid*., p. 89).

Because the residents did not trust the authorities, they sought outside assistance. Gibbs' brother-in-law, who was a biologist at State University of New York/Buffalo, assisted her and encouraged her to take a leading role. But even more important, he alerted another scientist, Beverly Paigen, who was doing cancer research at a nearby research institute. She conducted research on the relationship between toxics and cancer and became interested in the problem. At the time she tested out a hypothesis that some families exposed to low concentrations of toxics would be more susceptible to cancer than other families, and she regarded the Love Canal case as an opportunity to test her hypothesis.

Gibbs showed her map to the leader of the Department of Health's research group, but he showed little interest. Then she showed it to Beverly Paigen, who immediately became interested. Members of the home owners' association assisted her in interviewing 1140 of the residents of Love Canal. The result of the investigation showed a clear geographical distribution of ailments, and it appeared to follow a pattern that reminded of the earlier swales. Paigen then divided all houses into two categories: 'Wet homes' were houses built above or close to the swales, and 'dry homes' were all the other houses. The result was striking: Women in 'wet homes' had three times as many miscarriages as women in 'dry homes'. Birth defects, asthma, urinary infections and the number of psychiatric cases were several times higher in 'wet' areas than in 'dry' areas.

Paigen's investigation therefore supported Gibbs' original hypothesis. In addition it had clear consequences for what should be done. One should evacuate residents from the 'wet homes' first. This was contrary to the procedure selected by the health authorities of New York. They assumed that the toxics spread radially from the chemical waste site. Therefore, they had purchased the houses closest to the waste site and evacuated the residents who lived there, and planned to let the rest of the residents remain.

There was a confrontation. When Paigen's results were given to the press, the leader of the research group of the Department of Health said that the investigation was "totally incorrect", and other officials later argued that the evidence was based on "information collected by housewives that is useless" (quoted from Levine 1982, p. 93).

To make a long and dramatic story short: Beverly Paigen and the residents of Love Canal prevailed. In the summer of 1980 the Congress allocated additional funds that authorized the President to use up to 15 million US dollars to relocate the remaining residents. After a number of negotiations between different parties, Love Canal was for the most part vacated in 1981.

3. From facts to values

Beverly Paigen gave her version of the controversy in an article a few years later. She tells that she had originally believed that the case was a matter of scientific disagreement that could be resolved by having the involved researchers come together and compare the data, experimental design, and statistical analyses. But she was wrong. She was surprised to discover that the facts made little difference, and alleged that "it raised a series of questions that had more to do with values than science" (Paigen 1982, p. 29).

There were two main differences between the research group of the Department of Health and Beverly Paigen:

First, the research group had made the assumption that the toxics spread more or less homogeneously outwards from the waste site. This was in accordance with traditional scientific practice – going back to Galileo and Descartes – to start with the simple and idealized. It followed from this assumption that the area closest to the site contained the highest concentration of toxics, and the concentration would decrease with increasing distance from the site. Paigen, on the other hand, had adopted Gibbs hypothesis that the toxics dispersed along the swales.

Second, they held opposite views regarding the burden of proof and the uses of statistics. The research group had claimed that it had used a 'conservative' scientific approach. Paigen says that it occurred to her that there was a problem when she, in a conversation with a representative from the research group, discovered that they disagreed about how this should be interpreted in every single case they discussed. They both claimed to take a conservative approach, but it turned out that they had opposite opinions about what 'conservative' meant in this situation. The researcher from the Department of Health emphasized that one must be very cautious in concluding that Love Canal was an unsafe place to live. Paigen, on the other hand, maintained that one had to be very cautious about concluding that Love Canal was a safe place to live. She argued that since a mistake could result in severe consequences for the health of the residents, the researchers must be very careful in concluding that Love Canal was a safe place to live. She insisted that underestimating the danger was worse than groundless fear (*ibid*., p. 32).

Paigen's position is described in the article that I have quoted. Her approach can be expressed in a formulation given by Nicholas Ashford, director of Center for Public Policy Alternatives at MIT. The question should not be "can you publish this in New England Journal of Medicine, but would you let your daughter work with that chemical?" (quoted from Savan 1988, p. 59). We do not have written accounts of the position of the Department of Health scientists, for obvious reasons. When you apply traditional scientific methods and carry out what you and your colleagues regard as good science, you normally do not state this explicitly. However, we do have some quotations from conversations that support this view. Perhaps the best source is Adeline G. Levine's book *Love Canal: Science, Politics, and People*, which has been my main source of information on the Love Canal case. She followed the case closely from the very beginning and had close contact both with residents, politicians and scientists who were involved in the process. Among others she quotes "a high-level official" who emphasized that

> […] the health department professionals were *scientists*, who did not worry about people's reactions to cautionary statement and recommended actions. They dealt with numbers – with data on physical conditions – and only with these. Political and social matters, the official stressed, were extraneous to the DOH [Department of Health] work. [Levine 1982, p. 40]

Levine also pointed out that the scientists working for the Department of Health were afraid of losing their objectivity, and she quoted one scientist who explained objectivity in the following way: "We deal only with numbers; we are scientists" (quoted in Levine 1982, p. 85).

The positions of the Department of Health scientists and Paigen illustrate two of the categories introduced in Roger S. Pielke' book *The Honest Broker: Making Sense of Science in Policy and Politics* (2007). The book addresses the relationship between scientists and political issues, and it describes different ways scientists may regard their own role as scientists. One category is what Pielke calls 'Pure Scientists': Their role is to sum up the state of knowledge of a limited area, and leave it to the politicians to make decisions. The Department of Health scientists can no doubt be placed in this category. Two other categories are what Pielke calls 'Issue Advocates' and 'Honest Brokers'. They have in common that scientists belonging to these two categories take wider social and political concerns into consideration. The difference is that scientists belonging to the first category take sides in a controversy, and use their expert knowledge to pursue the political agenda that they support, whereas scientists belonging to the second category restrict themselves to pointing to the relationship between various options and political agendas. I think Paigen can be placed in the category of 'Issue Advocates', because she no doubt sided with the residents of Love Canal.

4. The Precautionary Principle: What it is and what it is not

Paigen's fundamental position – that because mistakes could result in severe consequences for the health of the residents, the researchers must be very careful in concluding that Love Canal was a safe place to live – was an application of what has become known as the precautionary principle (PP). I will in this section explain the principle and in the next section look at its preconditions and broader ethical and philosophical context. I will also discuss possible ethical justifications of the principle. Then I will apply it to the Love Canal case.

The account of the precautionary principle in this section follows the report by World Commission on the Ethics of Scientific Knowledge and Technology (COMEST): *The Precautionary Principle* (2005).

The principle, explicitly formulated, is of recent origin. It originated in the 1970s, but today it is best known by its formulation in the Rio Declaration:

> In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation. [Rio Declaration 1992, principle 15]

We see from the text that the term 'approach' is used instead of 'principle'. For our purpose it makes no difference, and I will not go into further details. Let us take a closer look at what the principle implies.

The text uses the expression 'lack of full scientific certainty'. This means that there must be some indications of possible damage. Ungrounded fear is not sufficient. It is also implied that scientific investigations have been made that make the assumption plausible. It is further required that the results of the scientific investigations are uncertain, and that this uncertainty cannot be reduced or eliminated before a decision is made. In the case of Love Canal both the research group of the Department of Health and Beverly Paigen had been involved, and there was indications of possible damage not only in the evacuated area, but in the surrounding area as well.

Second, the text refers to 'serious or irreversible damage'. This is open to interpretation. If a person dies, it is a tragedy for that person. And in larger accidents, many people may die or be injured. In an airplane crash hundreds of people may die. Therefore, one might argue that according to the precautionary principle we should ban aviation. But we do not ban all aviation because airplanes sometimes crash and people are killed. Uncertainty is a fundamental aspect of the human condition. However, the reduction of uncertainty is an important aspect of modernity, and the theories of probability

and statistics were developed as a part of modern science, with the aim of reducing uncertainty. The concept of risk was a part of this endeavor. Risk can be defined as the probability of an adverse quantifiable outcome (therefore, risk is sometimes defined as the product of probability and cost of damage), and risk calculations play an important part in decisions concerning choices of technology. On the one hand it is an acceptance of uncertainty, but on the other hand it is an attempt to control uncertainty (cf. the title of Ian Hacking's book *The Taming of Chance*, 1990).

Third, the text uses the expression 'cost-effective measures'. Again, this is open to interpretation. What is the value of good health, and what is the value of a human life, and what is the value of nature? One way of deciding the value of a harm is the compensation that is paid to the victims, or if the damage is reversible, the cost of cleaning up or repairing the damage. In many cases prevention is more cost-effective than cure, and therefore precaution is cost-effective. If precaution had been taken from the very beginning, the authorities would not have allowed the chemical waste to be buried in Love Canal, and if that had already been done, they would not have allowed the school and the suburb to be built on top of the waste site. However, this could not be undone. As the situation was now, the costs of cleaning up the place, or buying the houses of the residents and aiding them in relocating, paying medical bills, *etc.* would far exceed the costs of safely depositing the chemical waste in the first place. However, the problem is often that benefits and costs are not equally distributed. In Love Canal Hooker Electrochemical Company had got rid of the chemical waste in the cheapest possible way, and sold the land for one dollar on the condition that they should be exempt from any liability.

It was argued that even if Hooker did not have a legal obligation, it at least had a moral obligation. This was denied by the company. In the end the US Congress paid the costs because this had been a political issue. However, to decide in each case what 'cost-effctive measures' implies, is far from easy.

5. The Precautionary Principle: Broader context and justification

How can the precautionary principle be justified? I shall not try to give a complete justification, but restrict myself to elaborating some of the philosophical and ethical background that makes the principle plausible.

The precautionary principle may be regarded as a result of a change of scientific paradigm. Newton completed the scientific revolution that started with Galileo, and he made classical mechanics the model of all science. Be-

cause classical mechanics is deterministic, they established an ideal of exact predictions that was unprecedented. Using the laws of classical mechanics the motion of celestial bodies could be calculated thousands of years into the past and the future. However, at the end of the nineteenth century the mathematician Henri Poincaré showed that even a deterministic system may become unstable, making calculations difficult. In 1960 Edward Lorenz, who worked on a simplified weather model, showed that even in a simple deterministic system small errors may obstruct exact predictions (sensitive dependence on initial conditions, known as the 'butterfly effect'). His conclusion was that long-time weather forecasts are generally impossible. It was gradually acknowledged that the ideal of classical mechanics only applied to simple, idealized systems. Most natural systems are complex, and therefore uncertainty will in many cases be irreducible. The recognition that nature is complex and that uncertainty normally cannot be eliminated is an important precondition for the precautionary principle.

Another precondition is the recognition that nature is limited and vulnerable. To some extent this recognition represents a return to a view of nature that was dominating prior to the scientific revolution of the 17th century. The dominating view of nature from Antiquity through the Middle Ages was to regard nature as analogous to an organism (an organismic view), where man was regarded as a part of nature. However, this changed in the 17th century. The organismic view of nature was replaced by a mechanistic view, regarding nature as a big machine. The clockwork metaphor became important. Instead of regarding nature as our home, it was rather regarded as a resource that could be exploited for our benefits.

If we jump to our recent history, a highly influential book was Rachel Carson's *Silent Spring*. She brought the public attention to the widespread destruction of wildlife in America by the use of pesticides including insecticides. When the book was published in 1962, Carson was heavily attacked by the agricultural chemical industry and the scientific establishment. Questioning the benefits of the new pesticides was regarded as attacking the faith in scientific progress. However, president John F. Kennedy was so impressed by the book that he ordered the President's Science Advisory Committee to investigate the uses of pesticides in agriculture. When the committee published its report in 1963, it confirmed that Carson was basically right.

Carson's book was not only an attack on the uses of pesticides, but on a scientific and technological development that regarded nature as only a resource to be exploited by us. Although the process has been slow, it is today recognized that nature is both limited and complex.

Another factor is the recognition that we have a responsibility for nature and future generations. I will restrict myself to mentioning one influential book, by the German philosopher Hans Jonas. The book was first published

in 1979 in German under the title *Das Prinzip Verantwortung* (*The Principle of Responsibility*) and in English as *The Imperative of Responsibility* in 1984. As the original title indicates, Jonas made responsibility the very foundation of ethics, in opposition to the traditional view that to be a subject of ethics we require autonomy. His point of departure was that all ethical theories up until then had taken as frame of reference that all consequences of human actions are small in 'space and time'. However, the tremendous power of modern science and technology has changed this. The atomic bomb is the most dramatic example. A nuclear war will not only affect all people alive today, but future generations as well. It may even eradicate all life. Jonas therefore set out to develop the foundation of a global ethics, based on responsibility as the fundamental concept.

Responsibility is an asymmetrical concept: A may be responsible for B without B being responsible for A. For example, the parents are responsible for their children, but the children are not responsible for their parents. (When the children grow up and their parents are old, the relationship changes. But that is a different matter.) Therefore, the argument that we should not do anything for future generations because they have not done anything for us, is invalid. We may have responsibility for animals, future generations, and nature even if they do not have responsibility for us. Jonas was particularly concerned that we should not destroy the conditions of life for future generations. He even formulated a categorical imperative to replace Kant's categorical imperative. Like Kant he gave several formulations of his imperative. One of them is this: "Act so that the effects of your action are consistent with a continuing genuine life on earth." (Jonas 1984, p. 11)

At last I want to mention the article 'Asymmetries in Ethics', written by the Norwegian philosopher Knut-Erik Tranøy in 1967. It was written before the precautionary principle became a topic, and has hardly had any influence on the development of the principle. Nevertheless, in the article he introduced some important concepts that can be used to justify the principle. The article was inspired by the philosopher of science Karl Popper and his asymmetry between verification and falsification: Because scientific statements according to Popper are universal statements, they can never be verified. However, they can be definitely falsified. His favorite example was the statement: 'All swans are white.' Observation of thousands of white swans cannot verify the statement, but only one black swan can falsify it. Tranøy applies this asymmetry to some fundamental ethical concepts, like life and death, pleasure and pain, happiness and unhappiness, right and wrong, and good and bad. According to Tranøy these pairs of concepts are asymmetric, similar to verification and falsification in the sense that the negative terms are more definitive, categorical, and fundamental than the positive.

His first example is the asymmetry between life and death. We know numerous sufficient conditions for death, but whereas we know many necessary conditions for life, we do not know any sufficient conditions. Similarly, there are numerous ways of killing people, but to keep them alive, grow, and flourish is much more indefinite and vague. Therefore, it is easier to find a negative formulation, for example 'you shall not kill', than a positive. Although he does not deal with the pair utility/harm, it is obvious that the same asymmetry applies to this pair as well. Therefore, it follows that it is more important to prevent harm than to promote utility. This is a kind of 'negative utilitarianism' that can be used to justify the precautionary principle.

One might apply Tranøy's asymmetry directly to the Love Canal case. Utility would be the money saved if nothing was done. This money could be used for other purposes, and would increase the total utility in the New York State or in the country as a whole (depending on who actually supplied the money, New York State or US government). The harm would be possible health problems among the citizens of Love Canal. Of course, many good things might be done with the money saved, for example building new schools, or nursing homes for elderly people. But the money might also be wasted, on badly planned projects, corruption, or on an increasing bureaucracy. Although some uncertainty was involved, the harm to the citizens was rather concrete, causing severe health problems. Of course, uncertainty complicates the problem. Let us, therefore, return to the question of uncertainty.

6. The precautionary principle and statistics

Although the precautionary principle was not formulated at the time of the Love Canal controversy, it is interesting to see how it applies to that case.

There was general agreement that the inner, evacuated area was not safe. But what about the remaining part of Love Canal? The researchers working for the Department of Health concluded that it was safe, and their main argument was that the frequency of health problems, for example miscarriages, was not significantly higher than in comparable areas. Therefore, there was no causal relationship between the toxics at the waste site and health problems in the population of Love Canal (except the inner area that had already been evacuated).

One of the standard methods for establishing causal relationships is the use of significance tests. We compare one group which has been exposed to the alleged cause (the experimental group) with a similar group which has not been exposed to the same cause (control group). If we observe a difference in the two groups, we infer that it is due to the alleged cause. However, this

inference is only valid if we know that the two groups prior to the exposure to the alleged cause are similar (in all relevant aspects). Therefore laboratory conditions are preferred, because one can control the environment. But even under laboratory conditions it is impossible to control all factors. In particular, in all biological systems there is variation. One technique is to try to distribute the unknown factors equally in the experimental and the control group (randomization). However, all uncertainty cannot be eliminated, and outside the laboratory we have much less control.

Therefore, when we observe a difference between the experimental group and the control group, the difference may be a sign of a causal relationship, but it may also be due to factors beyond our control, or due to chance. But how do we know if it is the one or the other? One standard procedure is using a 'null hypothesis': We assume that the observed difference between the two groups has come about by chance. Only if the probability of this happening is lower than a certain level, called the significance level, will we reject the null hypothesis. A significance level of 0.05 (5%) is normally used. That means that only if the probability of the difference occurring by chance is less than 5%, will we conclude that the difference is not likely due to chance and that therefore we should seek another explanation – for example that the relationship is in fact causal.

7. Type I and type II statistical errors

We may make two kinds of error due to the statistical nature of a problem: A type I error is rejecting a null hypothesis when it is in fact true. This is equivalent to claiming that there is an effect that in reality does not exist ('false positive'). A type II error is failing to reject a null hypothesis that is in fact false. This is equivalent to overlooking an effect that really exists ('false negative').

Although the relationship between a type I and a type II error depends on the specific problem, in general there is a trade-off between them: If we decrease the probability of one, the probability of the other increases. In traditional significance tests the probability of a type I error is controlled (it is set by the significance level), thus leaving the probability of a type II error open. If researchers inform us that they have not found a statistically significant difference, it would be important additional information if the probability of overlooking the difference is large. In statistics text books it is therefore emphasized that the probability for a type II error should always be evaluated. However in practice this is different. We know that this rule is often violated.

One may ask why a type I error is worse than a type II error. One sensible answer is that scientists should have good evidence for asserting something, and when they have not, they should refrain from asserting. It is therefore more detrimental to their reputation to assert something that is not the case than overlooking something that is the case. Sometimes a legal analogy is used to justify the asymmetry. In a criminal case the jury sticks to the 'null hypothesis' of not guilty until the defendant is proved guilty. The burden of proof is on the prosecutor. The justification is that it is worse to find an innocent person guilty than failing to convict a guilty person (Bhattacharyya & Johnson 1977, p. 168).

To put the burden of proof on the one who asserts that there is a difference may make sense in basic science, but it turns out differently when used in applied science. (I shall later show that it does not make sense in basic science, either.) Engineers have treated type I and type II errors differently, because they know that their information is used to make decisions. Type I error is called 'producer risk' because a false positive may harm the producer: He may have to withdraw a product that is not really harmful, or pay compensation for a non-existing harm. Type II error is called 'consumer risk' because failing to detect a harmful effect harms the consumer. The traditional approach of minimizing type I error favors the producer and puts the burden of proof on the consumer or victim of pollution. For example, in environmental questions it is often the case that those affected by the pollution have the burden of proof, while those who pollute have the benefit of doubt. When Beverly Paigen and the residents of Love Canal claimed that the toxic substances from the dump site were the cause of illness, they had the burden of proof.

The scientists working for the Department of Health certainly knew who had hired them. One might suspect that they wanted to come up with results that favored their employer. This may often be the case. As a member of the Norwegian Research Ethics Committee for Science and Technology in the 1990s I attended meetings with researchers who confessed that they had conscientious problems because they knew that if they did not produce the results their contractor expected, future projects would be endangered. Needless to say, these confessions made a deep impression on the members of the committee. However, in Love Canal the scientists might just stick to traditional scientific methods, allegedly being objective and neutral. Therefore, insisting on this kind of method favors one part of the conflict, the 'producer' (or the one who is responsible for the harm).

8. Reversing the burden of proof

The Department of Health's research group claimed to do good science, whereas Paigen was regarded as mixing her personal feelings or political agenda into her scientific activity. As we have seen, she used precaution to justify her approach. Today we would just have referred to the precautionary principle to support her claims about the burden of proof, and one might argue that asymmetry between type I and type II errors should be reversed. In cases like Love Canal, where possible harm is imminent, it is certainly better to overestimate the danger than to underestimate it. The philosopher Kristin Shrader-Frechette argued in favor of this position as early as 1991 in her book *Risk and Rationality* (Shrader-Frechette 1991, Lemons *et al*. 1997).

However, Paigen might have taken a more assertive approach. Instead of defending herself, she might have accused the researchers of the Department of Health of doing bad science. To see how, we only have to ask the simple question about significance tests: Why do we use a significance level of $p = 0.05$ (or even $p = 0.01$)? The simple answer is that it represents a difference of two (or three, respectively) standard deviations from the mean if we assume a normal distribution. However, why two standard deviations? It is no more than a rule of thumb. It can be traced back to Ronald Fisher's *The Design of Experiments* (1935), where it looks as if only significant results according to this rule are acceptable. The problem is that Fisher neglected advice from all the other leading statisticians of his time. His simplified version prevailed, though, and became the leading orthodoxy until today. But it has drawn increasing criticism during the last decade. (For the historical background and criticism of traditional significance tests, see Ziliak & McCloskey 2012.)

I said previously that a possible justification for minimizing type I errors in basic science was that it is more detrimental to a scientist's reputation to assert something that is not the case than failing to detect something that is the case. This may be true as a factual description, but it is nevertheless bad science. As Karl Popper pointed out long ago, good science is not characterized by advancing modest hypotheses, but by advancing *bold hypotheses*. This kind of precaution is not a scientific virtue. However, when hypotheses have been advanced, they should be subject to severe tests. Indeed, we shall try to falsify them. If they survive the tests, we should keep them. But only temporarily, because hypotheses can never be verified. However, they can be definitely falsified, and will then be replaced by better hypotheses. We do not have to subscribe to all of Popper's philosophy of science to endorse his view that creativity and boldness are imperative to scientific progress (see for example Popper 1981, p. 96.) We only have to keep in mind the importance of

originality in scientific work. This is why plagiarism is regarded as one of the most serious kinds of scientific misconduct.

Many researchers have criticized significance testing. In 2005 the statistician John Ioannidis published the article 'Why Most Published Research Findings Are False'. The article focuses on the relationship between type I and type II errors. He uses typical values from research fields (for example, in some fields the probability of a type II error may be up to 0.5, which implies that there is a 50% chance of failing to observe a real effect), and some simple calculations to underpin his conclusion: "It can be proven that most claimed research findings are false." (Ioannidis 2005)

This year the American Statistical Association (ASA) issued a statement concerning the uses of significance tests. The conclusion is worth quoting:

> Let's be clear. Nothing in the ASA statement is new. Statisticians and others have been sounding the alarm about these matters for decades, to little avail. We hoped that a statement from the world's largest professional association of statisticians would open a fresh discussion and draw renewed and vigorous attention to changing the practice of science with regards to the use of statistical inference. [Wasserstein & Lazar 2016]

9. Uncertainty and ignorance

It is important to bear in mind that the significance level is more or less arbitrary, while it would make more sense if it was estimated in each concrete case. Additionally, it is important to keep in mind that significance tests only concerns random errors, due to chance. In most situations systematic errors are much more important: experimental errors, measurements errors, sampling bias, wrong models. *etc.* Love Canal gives a good illustration of this.

All these factors are sources of uncertainty that is much more serious than the risk due to random statistical variation. At least we must make a distinction between three different kinds of uncertainty:

1. Risk: This is the kind of uncertainty that we deal with in probability theory and statistics. We know possible outcomes, and we know the probabilities of various outcomes.
2. Uncertainty ('known unknowns'): In this case we know the possible outcome, but we cannot assign probabilities to them.
3. Ignorance ('unknown unknowns'): In this case we do not even know the possible outcomes and so we cannot estimate their probabilities.

We have seen that the research group of the Department of Health used a model where they assumed that toxics spread homogeneously outwards from the waste site. They expected to find the highest concentrations of both tox-

ins and health problems in the houses closest to the site, and decreasing outwards. However, this was not confirmed. They also took the total population in Love Canal outside the evacuated area and compared it to a similar population. They did not find any significant difference, and this supported their view that the rest of Love Canal was a safe place to live.

This conclusion was, however, based on an erroneous model. The assumption that the toxics spread homogeneously outward from the waste site was wrong, and the hypothesis that it followed the swales, was right. Therefore, they did not see significant differences. However, when Paigen divided the homes into 'wet' and 'dry', she had no problems in establishing significant differences (for example in the number of miscarriages). An important lesson from Love Canal is, therefore, that statistics does not help if one uses a wrong model.

This is an example of ignorance: What actually happened, was not part of the models applied by the scientists.

10. When the model is wrong

Experts trained in a field have a tendency to apply the kinds of models that conform with their field. The following example has been taken from Brian Wynne's 'Uncertainty and Environmental Learning' (1992): In May of 1986 a cloud of radioactive material from the Chernobyl accident passed over Cumbria in North Wales. Heavy rains caused a large amount of radioactive cesium to fall over an area used to raise sheep. The authorities in charge assured everyone that there was no cause for concern, but in spite of this, six weeks after the rains a ban against selling meat from sheep that had grazed in the area was imposed because of the high levels of radioactivity found in the meat. Experts claimed however that the radioactivity would rapidly decrease, and that the ban would be lifted in a few weeks. Yet, even after six years the level of radioactivity was so high in some of the affected areas that restrictions had to be upheld.

How could the experts be so wrong? Their predictions were based on extrapolations from the behavior of cesium in alkaline clay soil to the acid peat soil of Cumbria. Measurements showed that the dispersion of cesium in these types of soil was fairly similar, and on that basis they assumed that cesium would sink so far down into the ground that after a short period of time there would be no problem. This was based on the assumption that the radiation would come from the cesium in the soil and would be absorbed by people or animals who happened to be in the area. Under this assumption it was the physical transport of cesium in the soil that was important. However,

this assumption was wrong. The sheep received cesium in their bodies through the grass they ate. The important question was therefore not how the cesium was dispersed throughout the soil but if it was absorbed into the vegetation. Here there proved to be a significant difference between alkaline clay soil and acid peat. In alkaline clay soil, cesium adsorbs on aluminum silicate so that it cannot move into the vegetation, whereas in peat it remains chemically mobile and can therefore be taken up by the vegetation. The experts did not consider these possibilities, and that was the cause of their mistaken predictions (Wynne 1992, 121).

Should not a model that takes into consideration for example chemical properties, have been used at the onset? The answer is, of course, yes. But to understand why the experts made such an apparently elementary error we have to take into consideration that they had been trained as physicists. Physicists are used to think in terms of physical transportation and radiation. Chemists are trained to think in terms of chemical reactions and chemical mobility. The problem is that it is not a part of professional training to learn about the limits of the models and methods of a field.

A serious obstacle to coming to terms with this problem is the fact that Thomas Kuhn's description of the scientific community is to a large extent valid. We do not have to accept the more controversial parts of Kuhn's theory in order to agree that scientists are trained within a 'paradigm'. Parts of the paradigm will be the tacit knowledge which is imperative to everyday scientific work. This kind of knowledge cannot be articulated as explicit rules. Kuhn himself uses Michael Polanyi's term "tacit knowledge" (Kuhn 1970, p. 187). When experts deal with situations which fit into their paradigm, this works fine. But when confronted with situations that are not so easy to accommodate to the expert's paradigm, it is a source of error. Because experts in the same field are trained within the same paradigm, they are usually blind to many of their own tacit assumptions. However, experts from other fields may immediately be aware of some of the tacit assumptions of the field. Therefore, in cases involving complexity and uncertainty it is imperative to draw on various kinds of expertise.

11. Conclusion: Lessons

First lesson: Scientists are not outside

In this article I have presented two opposite views of scientific objectivity and the scientists' role in giving policy advice in cases where there are differ-

ent interests involved. On the one hand the scientists working for the Department of Health argued that their role as scientists required that they produce objective knowledge, independent of who might benefit and who might lose from this information. They acted as 'Pure Scientists' according to Pielke's terminology. On the other hand, Beverly Paigen argued that scientists should choose sides, and she applied what is today known as the precautionary principle. She acted as an 'Issue Advocate' in Pielke' terminology.

As pointed out earlier, the researchers of the Department of Health were victims of an erroneous idea of scientific objectivity. This prevented them from cooperating with the residents of Love Canal. If they had entered into a dialogue with the residents, they would have learned about the swales. Lois Gibbs learned about the swales from older residents. When systems are complex, local knowledge may be more useful than mathematical models. The result of this alleged objectivity and neutrality was that the researchers sided with one party, the authorities.

Should the Department of Health researchers have sided with the residents of Love Canal, like Beverly Paigen? That was not required. However, they should have taken the residents seriously, and cooperated with them. The result would have been better science, and better science advice. The first lesson from Love Canal is that scientists must not remain outside a controversy.

Second lesson: Inform about uncertainty

We have seen that statistics was an issue in the Love Canal case. However, I have argued that other kinds of uncertainty are much more important. The following example applies to all kinds of uncertainty.

Let us imagine that a group of researchers is assigned the task of examining whether there is a difference in the incidence of illness between two groups, and they give the answer T0: 'We have not found any (statistically significant) difference.' However for the decision maker essential information is missing, because the answer may be interpreted in two different ways. Either T1: 'We have not found any difference, and we most likely would have found it if it existed' or T2: 'We have not found any difference, but we most likely would not have found it even if it existed.' Needless to say, T0 would carry much more weight if the correct interpretation was T1 than if it was T2. Therefore, it is a serious problem when researchers answer T0, the majority of politicians and others interpret it as T1, and the correct interpretation is T2. This was probably the case in Love Canal.

Even if the scientists of Department of Health regarded themselves as belonging to Pielke's category of 'Pure Scientists', not informing about un-

certainty would not only be bad science advice, but bad science. However, it may have serious implications when decisions are based on scientific advice.

Third lesson: Laypeople should be involved

In their book *Uncertainty and Quality in Science for Policy* (1990) Jerome Ravetz and Silvio Funtowicz argue that science has to enter into a 'post-normal' phase to adequately address problems where uncertainty and 'decision stakes' are high. In the book they develop a conceptual scheme to deal with the new challenges. I shall not go into technical details, but will just mention one aspect of 'post-normal' science which is relevant to my discussion: the uses of what they call 'extended peer communities'.

'Extended peer communities' imply an extension of the traditional scientific community to include non-experts as well. However, this does not mean that laypeople should invade the research laboratories and carry out research. It does mean, though, that laypeople should take part in discussions of priorities, evaluation of results, and policy debates.

One reason for including non-experts is that they are sometimes closer to the problem. In Love Canal we saw that the residents had local knowledge that the scientists could not possibly have. Lois Gibbs did not have any relevant formal education, she was not trained in the uses of mathematical models and statistics. However, when she sat down with a map of Love Canal and put pins on every home that had reported health problems, she saw that they formed a pattern, and had the idea that the health problems might be connected with the swales that cut through the area. She thus made a contribution that was much more important than any of the scientists of the Department of Health. When problems are complex, local knowledge is at least as important as mathematical models.

The contribution from laypeople may also be valuable for the opposite reason. In Love Canal the contribution of the residents was valuable because of their closeness to the problems, but laypeople may also be valuable because they have a distance. Experts are often caught in their models, they are victims of 'tunnel vision'. One way of revealing experts' hidden assumptions is to ask apparently stupid questions, which may be regarded as an extension of an important element in the Socratic tradition in philosophy. We know that it was part of Socrates' strategy to pretend that he was more ignorant than he actually was. For example, in the dialogue *Gorgias* he asks the expert in rhetoric, the Sophist Gorgias, what rhetoric is. He then shows that Gorgias' answer is insufficient, and proceeds to 'deeper' or 'wider' questions. Very often the dialogue ends up with the fundamental ethical questions of the right, the good, and justice.

An ethical reason for bringing in common people is that they are affected by the decisions which are made. The questions of global warming, the ozone layer, radioactive waste, and genetically modified food concerns everybody, experts as well as non-experts. These questions are too important to be left only to the experts.

Acknowledgement

I want to thank Tom Børsen and two anonymous reviewers for valuable comments and suggestions.

Further reading

The most comprehensive account of the controversy is Levine 1982. A huge collection of original documents concerning the Love Canal case are available at http://library.buffalo.edu/specialcollections/lovecanal/collections/ (accessed 30 March 2016). There is a website for the precautionary principle: http://www.precautionaryprinciple.eu/ (accessed 30 March 2016). From this website you can download among others COMEST 2005. A critical history of significance tests can be found in Ziliak & McCloskey 2012. Many examples of the abuse of mathematical models, in particular in geology, are given in Pilkey & Pilkey-Jarvis 2007. Two recent collections of relevant articles are Pereira & Funtowicz 2015 and Meisch *et al*. 2015.

References

Bhattacharyya, G.K. & Johnson, R.A.: 1977, *Statistical Concepts and Methods*, Hoboken: Wiley.
Carson, R.: 2000 [1962], *Silent Spring*, London: Penguin.
COMEST (World Commission on the Ethics of Scientific Knowledge and Technology): 2005, *The Precautionary Principle*, Paris: UNESCO.
Hacking, I.: 1990, *The Taming of Chance*, Cambridge: Cambridge University Press.
Ioannidis, J.: 2005, 'Why Most Published Research Findings Are False', PloSMedicine, **2** (8), 696-701.
Jonas, H.: 1984, *The Imperative of Responsibility*, Chicago: University of Chicago Press.
Kuhn, T.: 1970, 'Postscript – 1969', in: *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press, pp. 174-210.

Lemons, J.; Shrader-Frechette, K. & Cranor, C.: 1997, 'The precautionary principle: Scientific uncertainty and type I and type II errors', *Foundations of Science*, **2** (2), 207-236.

Levine, A.G.: 1982, *Love Canal: Science, Politics, and People*, Lexington: Lexington Books.

Meisch, S.; Lundershausen, J.; Bossert, L. & Rockoff, M. (eds.): 2015, *Ethics of Science in the Research for Sustainable Development*, Baden-Baden: Nomos.

Paigen, B.: 1982, 'Controversy at Love Canal', *The Hastings Center Report*, **12** (3), 29-37.

Pereira A.G. & Funtowicz, S. (eds.): 2015, *Science, Philosophy and Sustainability*, London: Routledge.

Pielke, R.S.: 2007, *The Honest Broker: Making Sense of Science in Policy and Politics*, Cambridge: Cambridge University Press.

Pilkey O.H. & Pilkey-Jarvis, L.: 2007, *Useless Arithmetic. Why Environmental Scientists Can't Predict the Future*, New York: Columbia University Press.

Popper, K.R.: 1981, 'The Rationality of Scientific Revolutions', in: I. Hacking (ed.): *Scientific Revolutions*, Oxford: Oxford University Press, pp. 80-106.

Rio Declaration: 1992, *United Nations Conference on Environment and Development* (UNCED), Rio de Janeiro, 3-14 June 1992 (online available at: http://www.unep.org/documents.multilingual/default.asp?documentid=78&articleid=1163, accessed 21 Nov. 2016).

Savan, B.: 1988, *Science under Siege: The myth of objectivity in scientific research*, Montreal: CBC Enterprises.

Shrader-Frechette, K.S.: 1991, *Risk and Rationality*, Berkeley: University of California Press.

Tranoy, K.E.: 1967, 'Asymmetries in ethics', *Inquiry*, **10**, 351-72.

Wasserstein, R. & Lazar, N.A.: 2016, 'The ASA's Statement on p-values: Context, Process, and purpose', *The American Statistician* (online accepted version http://www.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108, accessed 30.3.2016).

Wynne, B.: 1991, 'Uncertainty and environmental learning', *Global Environmental Change*, June, 111-127.

Ziliak, S.T. & McCloskey, D.N.: 2012, *The Cult of Statistical Significance. How the Standard Error Costs Us Jobs, Justice, and Lives*, Ann Arbor: University of Michigan Press.

Ragnar Fjelland:
Center for the Study of the Sciences and Humanities (SVT), University of Bergen, Norway; Ragnar.Fjelland@uib.no