Science and Public Policy, volume 21, number 3, June 1994, pages 173-183, Beech Tree Publishing, 10 Watford Close, Guildford, Surrey GU1 2EP, England.

Real-life experiments

Society as a laboratory: the social risks of experimental research

Wolfgang Krohn and Johannes Weyer

***There* is *an increasing tendency to test research processes outside the laboratory. This requires an accompanying increase in responsibility, yet there are many examples of the irresponsible use of real-life experiments. Using a number of actual* cases *this paper raises the social dilemma of such experimentation*.**

Professor Wolfgang Krohn and Priv-Doz Dr Johannes Weyer are in the Fakultät für Soziologie, Universität Bielefeld, Universitätsstrasse 25, D-33615 Bielefeld, Germany.

This paper was written in connection with several research projects in which 'experimental implementations' of various kinds were studied (see Herbold, Krohn and Weyer, 1991; Herbold, Krohn and Weyer, 1992; Asdonk, Bredeweg and Kowol, 1991; and Weyer, 1994).

IN MODERN SCIENCE, there is an increasing tendency to extend research processes and their related risks beyond the limits of the laboratory or other such institution, and directly into the wider society. To greater or lesser extents, these moves are experiments. This tendency is apparent in, for instance, nuclear energy technologies, some developments in ecology, the use of physiologically dangerous chemicals, the introduction of some medicines, and in some military technology.

A common denominator in such cases is that they cannot be legitimised in the name of research alone. In keeping with at least the German way of viewing them,[1] they are declared to be the 'implementation of verified knowledge', and are justified on the basis of non-scientific (for instance commercial) interests.

With the increased rate of technological innovation in recent decades, the implementation of verified knowledge has frequently become *(nolens volens)* the testing of risk-involving technologies. For want of a better term, we will refer to these processes as 'experimental implementation(s)', 'real-life experiments' or 'implicit experiments'. The 'risks' of interest here are related to safety, not to (for instance) the probability of obtaining publishable scientific results or of making profits.

These terms have been chosen to suggest that the introduction of new technologies frequently reveals features of the traditional laboratory experimental generation of knowledge. Indeed, this is often deliberately built into the design of the project. Real-life experiments, however, differ fundamentally in one aspect from traditional experiments: they overstep the institutional bounds usually established for experimental research. From another perspective, one could say the limits of experimental science are being extended.

For instance, whether or not the release of genetically manipulated bacteria is harmful to the environment is something that can be finally determined only by trying it out. When society takes such a role in science, it becomes an "experimental society".[2] It necessarily takes upon itself the scientific risks of some types of error and failure.

When the testing of knowledge is carried out in non-scientific institutions (for instance, in operational nuclear power plants), society at large is exposed to the dangers of scientific error. This requires a redistribution of responsibility.

Science has the problem of being unaccustomed to the dilemma of the responsibility for the risks emerging from such experiments.[3] Our focus is on an analysis of some aspects of this development; we will not go into what institutional or legal changes would be necessary to achieve this redistribution of responsibility.

The problem is in the relationship between safety theories and experience with implementation. Here, accidents can be tests of the theories.[4]

Interpretations of accidents as part of a scientific experiment face the objection that accidents cannot be elements of a deliberately planned acquisition of knowledge, as they are unexpected or at least unintended events.[5] However, the hypothetical, anticipatory description of risks linked with these technologies does handle the possible occurrence of such accidents partly by means of scientific simulation and prognosis, seeking to prevent them by technical means.

Such accidents are not caused by the hand of fate, but result from the taking of risks.[6] Still, it remains incontrovertible that accidents are not deliberately conducted as experiments. On the contrary, it is precisely for the purpose of avoiding them that theories of risk research and models of risk simulation have been developed.

Given that these theories whose verification requires observable events, their testing is dependent on exactly the accidents they serve to avoid, and, indeed, on any other accidents which may occur.

Theory of real-life experiment

A 'trial world'

Fundamental to the understanding of science in the sociology of science is the concept of an institutionalised, 'free' research arena for the production of knowledge. This arena is defined in terms of experiment and hypothesis; its social manifestations are the laboratory and scientific discourse. Among its constitutive elements, there is the assumption that theoretical statements and conclusions, as well as success and failure in experimentation, arc free from moral considerations. The exemption of scientific action from the social consequences of error and failure was gradually won during the 17th century in the form of the granting of royal privileges; today it is considered to be a fundamental right. Nonetheless, it applies only when two ideal conditions are met:

1. Operations in the 'containment' of the laboratory must be (practically) without effect outside the laboratory; this is to be ensured by an experimental design guaranteeing that the operations be reversible or interruptible, or that their effects be sufficiently small as to be negligible.
2. Theoretical claims made in the 'containment' of scientific discourse must similarly be (practically) without effect on everyday discourse outside science. Their validity depends on the conditions of acceptance within science; if such acceptance is not forthcoming or withdrawn after being granted, it must be possible in principle to terminate their circulation in society.

When these conditions are fulfilled two consequences follow, one of a sociological nature, the other having to do with the theory of knowledge. Science makes allowances for the production of mistakes and failures more than docs any other institution in society. Indeed, it is precisely this tolerance for mistakes that has led to the acceleration in the production of knowledge characteristic of the modern age.

Secondly, experimental-hypothetical science always deals with idealised objects and segments of reality. The feedback between theory and experiment leads to the formulation of models of reality whose existence is ensured under boundary conditions that can only be realised in the isolated world of the laboratory. No laboratory scientist would change a theory only because some instrument works erratically or the energy supply is cut off. But in the 'real' world of technological implementations occurrences of this kind can necessitate dramatic changes in a theoretical model.

Obviously, the 'isolation' of the world of experiments and theories is itself an idealised construct that can only be approximated. In fact, action within the laboratory always intrudes upon reality as irreversibly as any other action. In order for such action to satisfy condition (1), however, the damage incurred must be 'small'. Further, theories cannot be limited to scientific discourse, but leave irreversible traces on societal communication, often without any intra-scientific consensus concerning their reliability. In cases in which religious, political and cultural traditions are affected by theories (prominent instances have been: Darwin-

> The development of modern science has come about through the institutionalisation of a special, closed world of research institutions within which knowledge is produced; responsibility for its application rests with the institutions of society

ism, psychoanalysis, Marxism) condition (2} cannot obtain.

Although such cases of conflict arising either from a new idea or a new technology indicate the limits of the ideal construction of hypothetico-experimental science, in general the development of modern science has come about through the institutionalisation of a special, closed world of research institutions. Knowledge is produced within this world; responsibility for its application rests with the institutions of society. Pasteur expressed this traditional view in the claim: "There is no such thing as applied science; there is only science and its application".[7]

Application of knowledge

The philosophy of science has rarely dealt with the question of the application of scientific knowledge. Nonetheless, it provides a key for analysing the mechanism through which the processes of implementation become research experiments. The alternatives are simply stated: if scientific knowledge is to be 'applied' in society, then either the existing boundary conditions for laboratory science must be adapted to those of society, or societal practice must be changed according to the standards set for laboratory science.[8]

Traditionally, the technical application of knowledge has not been considered to be a phase of scientific knowledge production. However, the application outside science of knowledge that has been scientifically validated is always the production of new knowledge: the implementation of new technologies, while being the application of tested knowledge, is at the same time a means of obtaining new experience, that is knowledge that could not have been acquired otherwise.

Generally, disputes about the behaviour of new technologies remain inconclusive as long as the technical methods are not put to work. This is especially the case in the question of reliability: for example, the US space shuttle was launched several times, although its reliability for manned flight had not yet been confirmed, in order to save money on expensive pre-tests. These flights served in turn to justify lower reliability standards. The astronauts were protected during these experimental launches by ejector seats which were later removed.[9] Generalising her analyses of military technology R&D Kaldor states that controversies about the performance of sophisticated new weapons systems usually remain unsettled unless the systems are tested under realistic conditions.[10]

Several case studies have shown that the implementation of theoretical models under 'realistic' boundary conditions leads in turn to the construction of new models.[11] MacKenzie demonstrated in the case of missile accuracy that tests are performed, among other reasons, "to help produce, and crucially 'verify' a mathematical model of the causes of system inaccuracy; ... part of the point of the test procedure [was] to assess the model".[12]

In other cases only negative experiences following the implementation of techniques that had been presumed reliable have revealed a previously low degree of scientific accuracy in the modelling of the circumstances and conditions. While the Aswan High Dam was seen at the time of its construction at best as an example of a complex technology, and not as an experiment with the ecolog) of a river basin and part of the Mediterranean Sea, it is now generally accepted that installations of this kind are complex experiments with the environment, the climate and tectonics.[13]

So far, the term 'experiment' has been used without being defined. It is now time to explain its use in the present context. On the one hand, to call 'experimental' virtually every kind of societal change would be too extensive; on the other, restricting the term to laboratory research would, by definition, mean excluding what we have called real-life experiments. We would like to characterise the notion of 'experimental research' as follows:

- the drawing up of a theoretical design and the formulation of hypotheses concerning the occurrence of (future) events;
- the setting up of an experimental situation that can serve as a means of testing theoretical assumptions; this includes knowledge of the relevant boundary conditions, the controlled variation of parameters as well as the making of arrangements for the observation of effects;
- the establishment of an organised research process embedded in a network of scientific institutions.

When these three conditions are fulfilled in a coordinated manner, an experimental situation is present, regardless of whether all three factors come together in one action strategy. This means that non-scientific action can also be part of an experiment.

Ideally, of course, all three conditions can be best realised in an organised research process in the laboratory, whereas any overstepping of the bounds of the laboratory necessarily results in the (partial or total) loss of control over the boundary conditions and in a limiting of the admissible variation of the test parameters. Nevertheless, real-life experiments also presuppose the existence of an organised research process; otherwise, it would be more appropriate to speak of hazardous evolutionary changes.

There are, however, intermediate cases in which, through the help of so-called 'monitoring sciences', certain dangerous developments become reinterpreted as risk-involving experiments. By describing such developments as risk-involving, scientists try to awaken a sense of responsibility in other actors for the possible consequences.

On the other hand, science finds itself burdened with responsibilities it is not accustomed to. One possible consequence here is that scientists' theory-based assertions about the risks of technological implementation could themselves become the subject of legal cease and desist actions. Thus risk perception magnifies the tendency to conceive of technical innovation as experimental testing.

Patterns of experimental design

In this section we shall discuss several cases of experimental implementation that illustrate different ways in which science oversteps the limits of the laboratory. We distinguish four categories of real-life experiments, based on differences in the underlying cognitive interests and research questions involved. Furthermore, we will show that the participation of scientists always occurs within complex networks of actors whose behaviour is only partly controllable by science. While this has serious consequences for research, it provides opportunities for carrying out experiments that would otherwise not be possible for science.

Our main goal in analysing the following cases is the elucidation of existing research processes and their corresponding experimental designs. These cases may appear rather strange for the sociology of science, whose analytical tools are not well adapted to the description of knowledge production outside the laboratory. We, too, cannot do more than direct our attention to certain characteristic, artificially induced effects, which by their very nature can only partially be studied under laboratory conditions.

Accidents in complex installations

As already mentioned, accidents in complex installations are rarely completely unexpected occurrences. Their probability and their possible course have been modelled theoretically by means of a body of knowledge gathered by research on safety problems since the days of the steam engine. The main reason why this research cannot be confined to the laboratory is that no security theory can take account of the risk involved in not taking certain risks into consideration. In the early days of risk analysis, Clifford Beck, deputy director at the US Atomic Energy Commission (AEC) and head of a steering committee for revising the first AEC safety report (WASH-750 from 1957), wrote in an internal memorandum (1965):

> "There is not even in principle an objective and quantitative method of calculating the probability of improbability of accidents or the likelihood that potential hazard will or will not be taken."[14]

Beck argued that, without assuming certain conditions for the occurrence of accidents, no calculation is possible, which means that, in principle, the procedure for selecting these conditions cannot be calculated. This scepticism appears only in toned-down form in the report, which, moreover, was only published in 1973 in response to a Freedom of Information Act request.

Nevertheless, nuclear power plants turned out to be a 'good' case in point. In 1975 the Browns Ferry reactor suffered a serious fire caused by the absolutely unanticipated handling of a candle. The Reactor Safety Study of 1974 ('Rasmussen Report') did not even consider cable insulation fires as an initiating event for a major accident, let alone the possible consequences of using a candle for tracking down gas leaks. In a memorandum of the US Nuclear Regulatory Commission of 1977 reference is made to simulation tests of the electrical security systems performed under the new assumption that candle-light might be dangerous, which yielded unexpected deficiencies in almost all American reactors.

Ironically, in countries where public discussion of risks is suppressed scientists can state more openly that the quality of knowledge about security depends on experience. The following appeared in a study by experts from the former GDR (German Democratic Republic):

> "Reliable assessments of the overall risks of nuclear power plants are not to be found in the literature. To determine this risk remains a rewarding goal of nuclear technology. It is only through many years of careful collection and evaluation of experience in existing nuclear plants that it will be possible to approach this goal."[15]

The dilemma resulting from learning by means of accidents that have to be prevented as much as possible in advance can be illustrated by many cases. We choose a recent example from air traffic. This is a highly reliable everyday technology, exposed to an equally high level of public attention, leading in turn to a permanent demand to reduce the risks still existing by implementation of safety and redundancy components on board and in the air traffic control centres.[16]

Although air traffic is highly reliable, the mechanism of learning about the functioning of the system by means of accidents cannot be eliminated. On 8 January 1989 a Boeing 737-400 of British Midland Airways crashed, killing 44 people.[17] The main cause for the accident was an "unforeseeable" coupling of the failure of three components, two of which were safety components.

Firstly, one of the CFM-56-3 engines of the twinjet began to burn. This normally does not raise serious problems, since the engine can be shut down, the fire automatically extinguished and the plane flown using only one engine to the nearest airport.

Secondly, the fire-warning system is thought to have been miswired so that the pilot unknowingly shut down the wrong engine. The burning engine then exploded in the final stage of the airport approach (when it had to perform at maximum power), which finally caused the crash.

Thirdly, the crew apparently did not double-check the warning, by which they could have discovered that a false signal had been given. It is scarcely possible to train someone under realistic conditions to deal with such situations. "Flight crews and training officials have had difficulty in constructing a scenario which would fit the known facts."[18]

The facts known at the time were: the engine CFM-56-3, which had been built 2,800 times by the end of 1988, was known to be problematic even before the accident. Several incidents had occurred in 1987 and 1988,[19] and the FAA (Federal Aviation Administration) had recommended that CFM-56-3-powered planes not be flown at high humidity and low temperatures.

The practice of Hying the Boeing can be reconstructed as an implicit real-life experiment, designed (a) to test the technical system under realistic conditions and presumably beyond the limits set by the authorities at the FAA (such statistical effects as the overall in-flight shutdown rate of the engine could be determined); and (b) to examine hypotheses concerning safety and emergency measures especially at the man-machine-interface.

Closely related to the category of learning by means of accidents, but not necessarily associated with the modelling of risk and safety, is the testing of complex technologies under realistic conditions

Improvement of prototypes

Closely related to the category of learning by means of accidents, but not necessarily associated with the modelling of risk and safety, is the testing of complex technologies under realistic conditions. The cases in this category cither involve unique models ('unica'), which are improved little by little, or a series of production types, which is improved from type to type, or technologies somewhere between these extremes, such as pipelines and dams.

Once again, nuclear power plants figure prominently here. In Germany both the HTR (high temperature reactor) in Hamm-Uentrop and the fast breeder in Kalkar are occasionally (and correctly) referred to as "research installations".[20] The HTR (no longer operational) was designed for investigation of the concept of 'intrinsic security', while the Kalkar reactor (never completed) was to be used to study the plutonium circuit.

We present an example from the military. Here it is well known that many high-performance technical systems can only be tested under actual combat conditions, that is in situations in which they are supposed to perform at maximum level, but which can neither be anticipated nor simulated. For example, it was clear in advance to many experts that SDI (strategic defense initiative) software could not be written faultlessly, which meant that the system's first 'real' application could well result in a catastrophe.[21]

Our concrete example will be a 'mini-SDI' named Aegis, sometimes called "Star Wars at Sea". The system is installed on several ships of the US Navy. According to its constructors Aegis has the capability of defending ships for hours against simultaneous attack by several warships and submarines and hundreds of aircraft and missiles.[22]

During the first demonstration of the software the system failed to hit six targets out of 16.[23] Since then, testing in several real combat situations has both confirmed the critics of SDI and demonstrated the social dilemma of real-life experiments. Several hundred people were killed due to Aegis' malfunctioning in 1987 and 1988. The battles between warships of the USA and of Libya in March 1986 provided the first opportunity for operating Aegis under combat conditions.[24]

It turned out that identification of Allied missiles, which were used by Libya, had not been programmed into the Aegis software, since the system had been constructed for a confrontation with Soviet technology.[25] In another instance, on 17 May 1987, the USS *Stark* was hit in the Persian Gulf by two Exocet missiles fired from an Iraqi aircraft; 37 seamen were killed in an accident which the electronic surveillance system AN/SLQ-32 was designed to prevent.[26]

The Armed Services Committee's report confirmed that the Iraqi aircraft had been discovered by the USS *Stark* more than an hour before the attack, but had not been taken seriously because Iraq was not classified as a hostile nation. The Exocet missiles, which could easily have been destroyed by the *Stark's* defense systems, were not identified because the ship had not "been turned broadside to the approaching Mirage"[27] and the ship's superstructure had created a blind spot which blocked the radar.

On 4 July 1988 Aegis failed once again, when the USS *Vincennes* shot down an Iranian Airbus, killing the 290 people on board, believing it to be an F-14 attacking the ship. The subsequent investigations "confirmed" that the system had worked properly and shifted the blame onto human error.[28] As usual, this kind of interpretation is on shaky ground, since improper construction of the machine-man interface is a human error by the designers, who do not have sufficiently reliable data about the coupling of human activities and machine behaviour under realistic conditions.[29]

Further, the data tapes of the *Vincennes* revealed that an ascending and not a descending plane had been tracked. But the combat information centre had reported a descending plane with certain military characteristics, which led to the captain's decision to shoot it down. The question was later raised of "whether the system could process the information fast enough to provide an in-depth analysis of... the aircraft".[30]

A further question raised was why a high-tech radar able to distinguish incoming objects on a scale of some tens of centimetres had been unable to distinguish a 54-metre airbus from a 19-metre F-14.[31] A more detailed investigation brought to light a partial breakdown of Aegis some days earlier.[32] The result had been that Aegis was not able to provide proper information about the characteristics of a single object over a period of eight minutes. In the case of the USS *Stark* it had not been able to provide protection against objects coming in at specific angles. This calls into question its alleged capability of simultaneously fighting against several hundred objects.

How well Aegis will function even in such limited war scenarios (which are child's play compared to operations in a full-scale war) can only be determined by means of implicit experiments in which the system's designers can gain the necessary experience.[33] The real-life experiments revealed an unexpectedly narrow realm for the employment of Aegis: neither a confrontation with Allied weapons, nor combat in a restricted area such as the Persian Gulf, were part of the system's design.[34]

Long-term cumulative effects

So far we have discussed cases that have to do with security theories and the reliability of technological implementation. The analysis of cumulative effects usually does not deal with complex technologies as such, but with artificial substances. It focuses on two issues: first, the physiological and/or psychic effects of (usually) small quantities of a substance, or low doses of radiation, which human beings are exposed to over a long period of time; and second the damage to natural resources caused by the long-term accumulation of various substances.

In many cases there is an overlapping of the two situations. In the early and famous example of the exposure of forage, fruits and vegetables to the insecticide DDT, scientific research gradually discovered that

- DDT accumulated in the fat tissues of female animals and women and was transmitted to newborn children;
- mutant DDT-resistant strains of insects developed; and
- the insecticide accumulated in soil and groundwater.[35]

In 1948, when the career of DDT had just started, the American Medical Association issued the warning that its chronic effects were "entirely unexplored".[36] A further warning from 1960 said:

> "It is a striking fact that knowledge of its mode of action has rarely preceded the use of any insecticide. Even today we do not know precisely how DDT induces its toxic action."[37]

Since the effects to be studied arc not properties of the substances themselves, but properties that emerge in their interactions with complex environments, they can only be investigated insufficiently in the laboratory. Further, it is not certain what one is to look for, nor for how long. In addition, the samples needed for a broad, long-term epidemiological analysis cannot be dealt with in the confines of the laboratory.

Another example reveals the virtually criminal nature of certain real-life experiments. In 1986 the Subcommittee on Energy Conservation and Power of the US House of Representatives released a report entitled *American Guinea Pigs: Three Decades of Radiation Experiments on U.S. Citizens.*

The report describes a series of partly coordinated experiments which had been conducted for the purpose of gathering knowledge about the effects of different levels of radiation on human beings. Most of them are 'classical' laboratory experiments which would have led to prosecution because they violated both specific laws and human rights. But some are complex real-life experiments: from May 1963 to November 1965 several experiments funded by the Atomic Energy Commission took place at the National Reactor Testing Station in Idaho, to

> "improve knowledge of the transport of radioactive iodine ... through the air-vegetation-cow-milk sequence in the human food chain. This information was necessary for developing reactor siting criteria ... and as an aid in planning emergency measures to be taken in the event of a radiation accident."[38]

Humans were exposed to radiation either by inhalation on the well-monitored test pasture or by drinking milk from cows which had grazed for several days on this so-called "hot pasture". The details were analysed by scientists and the experimental design was frequently changed in subsequent tests to measure different effects.[39]

In other experiments in 1956 US Air Force planes were sent through radiation clouds.[40] In 1963 the Columbia River was deliberately irradiated by the Hanford nuclear plant, in order to develop a "method for measuring the body burden caused by low-level radionuclides in humans".[41] People who had eaten fish from the river were examined.

The report of the Committee reveals that these tests were scientific experiments performed on the environment and on human health. They had an explicit research design and were monitored by scientific devices for data collection. They were carried out in order to explore the phenomenon of radiation, that was a new field for science, as well as for civilian and military technology. They were part of a broad research programme in nuclear physiology carried out in the US by institutions such as MIT (Massachusetts Institute of Technology), Los Alamos, the University of Rochester, Massachusetts General Hospital, and the University of Chicago.[42]

Another long-term research programme carried out in the open field of non-scientific institutions involved dioxin.[43] Dioxin was discovered as an (unwanted) by-product in herbicides and wood preservatives based on 2,4,5-trichlorphenol. The substances entered the market around 1950 and were manufactured by Boehringer (Germany), Dow (USA), Monsanto (USA), as well as other firms.

Unanticipated mass outbreaks of skin disorders ('chloracne') among workers were observed in 1949 after a chemical accident at Monsanto in West-Virginia and under normal production conditions in Hamburg at Boehringer.[44] A scientist at the University of Hamburg who collaborated with Boehringer identified the poisonous substance as 2,3,7,8-tetrachlorodibenzo-p-dioxin. It turned out to be harmful when brushed on the ear of a rabbit in solutions with a concentration of only 0.001%, making it one of the most poisonous substances ever discovered.

The chemical engineers at Boehringer altered the production process and achieved a considerable reduction of the by-product. The scientific results were published in the leading journal *Chemical Abstracts* in 1958, and in 1957 Boehringer informed all manufacturers of these herbicides world-wide both of the dangers and the newly developed production method.

> Although laboratory experiments and controlled field tests can bring to light some of the potential dangers, it is an illusion to assume that they can be used to investigate all the relevant factors sufficiently broadly and long enough

This case represents normal experimental practice in the chemical industry. The development and release of thousands of new substances a year does not allow for a monitoring system for the complete prevention of potential hazards. It is precisely because the chemical industry is science-based that the investigation of long-term effects and the possible side-effects of substances cannot precede their use in production and consumption processes by which these very effects are generated.

Although laboratory experiments and controlled field tests can, of course, bring to light some of the potential dangers, it would be an illusion — one much fed by governmental and industrial propaganda — to assume that such tests can be used to investigate all the relevant factors sufficiently broadly and long enough, and that the statistical models used for extrapolations are reliable.[45]

But research on dioxin was only in its infancy in 1957 and the rest of the story reveals how much can still be learned, even when the dangers are basically known. The letter from Boehringer was 'filed and forgotten' at Dow. Owing to the world-wide demand for herbicides, production based on the uncorrected procedure was increased everywhere. In 1961 it was discovered that the workers also suffered from psychological disorders. In 1962 the first reports of customers experiencing chloracne came in.

In 1967 "another kind of... familiar unplanned worker experiment"[46] was performed when a chemical reactor exploded in Grenoble. The use of the word "familiar" is justified, because similar accidents occurred in the Netherlands (1963), in Derbyshire, England (1968) and finally in Seveso, Italy (1976).

However, the real large-scale experiment took place during the Vietnam war, lasting for a decade (1961-1970). "Vietnam was used as a testing ground for chemical and biological warfare".[47] 'Agent Orange' was a mixture of the herbicides 2,4-D and 2,4,5-T. The evaluation of data collected by American experts during the war brought to light that the herbicides were an environmental hazard and also caused birth defects. In 1970 Samuel Epstein, who had been chairman of a federal panel on the teratogenicity of pesticides in 1969, declared: "Continued use of these herbicides in the environment constitutes a large-scale human experiment in teratogenicity".[48] This was intended as a warning, but it was, in fact, a description.

In the subsequent years Dow had several lawsuits brought against it. In a trial brief it was claimed that Dow "had not done research on the long-term effects of exposure".[49] In reality, however, the exposure *was* the research process. Looking back in 1983, Ted Doan, president of Dow in the early years of the Vietnam War, declared his philosophy:

> "Everything is poisonous. But you cannot prove that something isn't harmful. It may be that one molecule of dioxin ... can cause cancer ... The chances of this being true are almost zero. But you can't prove it isn't true."[50]

If this was an argument for continuing the research programme in dioxin, the social conditions for its acceptability have since changed.

Non-linear and recursive effects

Recursive dynamics has become an important field of study in many scientific disciplines. It plays an important role not only in laboratory research in such areas as laser optics, thermodynamics and solid-state physics, but also, for instance, in modelling the dynamics of the weather, the climate and the environment.[51] A strong tenet of the theory is that even if a recursive system is deterministic its future development cannot be predicted. The classical hierarchy of theory and experience no longer applies, for the theory states that one cannot predict what will be observed. The result is that experimental interest is *theoretically* directed towards the discovery of surprises.

Since it is believed that many natural systems are recursive, a good part of today's interventions in nature are experiments with recursive systems. If this is so, then, by definition, little can be done to anticipate the outcomes and dangers of such experiments. Among the frequently discussed examples are the deforestation of the tropical forests in Brazil, Thailand and the Congo, and the destruction of the ozone layer. In both cases scientific hypotheses concerning the effects on the global climate (in the case of the ozone layer) and on living beings have been, and are being, developed, and exhaustive data on these world-wide experiments is being collected.

The research project Airborne Antarctic Ozone Experiment has been in existence since 1987, with the participation of 150 groups of scientists throughout the world. A commentary on it stated:

> "There is at least one positive aspect to the ozone hole. In addition to convincing the international community to cooperate in reducing a threat to the environment, the hole has spurred investigators to study atmospheric chemistry and dynamics in greater detail. This effort has revolutionised our knowledge of how ozone interacts with other gases and how the interactions arc affected by meteorological conditions."[52]

Another example is the hypothesis of nuclear winter as an effect of a nuclear war. It was developed by John W Birks from the University of Colorado and Paul Crutzen from the Max-Planck Institute in Mainz in 1982 and provides a typical case for the creation of risk-perception by scientific research.

The theory hypothesises a drop in global temperatures of roughly 20°C due to the smoke from fires and burning caused by explosions.[53] Other theorists developed the counter-hypothesis of an increase in the ground temperature due to a warming-up effect on the smoke-clouds by the sun. What will really happen cannot be reliably predicted, because recursive processes are involved. In 1987 an

> "unexpected natural trial [of the hypothesis] occurred, when a severe thunderstorm set fire to a forest area of 203 square kilometers at the border between California and Oregon. This is about 0.1 percent of the area it is supposed could burn in a world-wide nuclear war. The meteorologists used the opportunity to compare the data obtained with climate data from previous years and in neighboring areas."[54]

The data showed a drop in temperature of 20 C as compared with unaffected areas. More detailed analysis revealed an important recursive mechanism by which the cooling-down process fed itself.

It may be doubtful that such a natural experiment should be presented among the cases described above. While we in principle do not see a strict dividing-line between artificially induced experiments with natural systems and natural disturbances of such systems, our interest in mentioning the case lies in its connection with the initial hypothesis of nuclear winter: the natural occurrence is taken to be a small-scale model of what would happen if the artificially induced experiment were to take place.

> "This local cooling-down cannot be taken as proof beyond any doubt ... It does show, however, the correctness of the basic assumptions for simulation of the impact on the climate of a nuclear exchange and that the order of magnitude can be taken to be realistic."[55]

Concluding remarks

The cases discussed under the four categories above lead to five general conclusions:

1. In a number of areas of research the production of scientific knowledge takes place outside the laboratory. This is because the mode of operation of complex technologies, particularly as regards their interaction with the real environment, cannot be studied within the confines of the laboratory (example: Aegis), or the risks entailed by interaction with the environment can only be investigated by means of implementation on the scale of the real world (example: dioxin). The risks resulting from the potentially fatal malfunctioning of technical systems (example: Boeing) or from the anticipated burden placed upon the ecosystem and on human health (example: the ozone hole) define situations of 'relevant lack of knowledge', opening up fields for real-life experiments.
2. There are various research strategies in which the structural pattern of real-life experimentation is employed. The universality and flexibility of the experimental method within the laboratory are expanded still further when it is used to gain knowledge outside the laboratory. Although this expansion is subject to new kinds of restrictions arising from the size of the objects, the magnitude of the risks, the irreversibility of the situations, and so on, there is nevertheless a broadening of the strategics of acquisition of knowledge by experiment.
3. In all the cases discussed above it was shown that research designs were present. In some cases they were explicit and publicly known (example: the ozone hole); in others it was not until an accident occurred that the experimental character of the technology was revealed (example: Boeing); in yet other cases, a criminal investigation was necessary to bring the research processes to light (example: radiation experiments). Although in some cases irresponsible and even criminal behaviour may play a role, this is only a secondary aspect of our argument.

   The exclusion of all manifestly unlawful research would in no way eliminate the social dilemma of experiments in the real world. The fact that 'under-cover research' takes place also shows that neither scientists nor the public nor the legal system are sufficiently prepared to face the fact that real-life experimentation in and with society has become an unavoidable part of social and scientific development.
4. For systematic reasons, there are parts of scientific research that cannot be consequence-free. Whole branches of research would have to be dismantled were research to relinquish its element of societal action. This incorporation into society is revealed in the 'coincidence between research and implementation'.

   When research is faced with a lack of experimental confirmation or the non-reproducibility of technical, social, human or ecological complexities in the laboratory, it either becomes dependent on implementations for the validation of knowledge, or implementations lead to new kinds of research questions, providing a primary empirical basis for the production of new knowledge. In both instances, research operates in the context of non-scientific practice.

   The cognitive challenge presented by such a production of knowledge lies in the fact that a stipulation of the boundary conditions based solely on theoretical considerations would not be acceptable to the public. Making human failure, unforeseeable disturbances or unknown side-effects responsible loses, in principle, its legitimacy, and shows up instead a problematic weakness in the disciplines involved. These disciplines can no longer afford to be indifferent to what happens 'outside' the bounds of theory. This forces research to integrate the boundary conditions as system variables and to break out of its discipline-bound restrictions.
5. With the development of society towards an 'experimental society', it becomes increasingly involved in negotiations concerning the conditions for the performance of experiments in and on society.

Notes

1. For one legal foundation for, and restrictions on, research see the commentary on the West German *Grundgesetz* (Constitution) by Maunz et al (1983, Rdnr 81ff, Art 5 III).
2. This concept is related to Beck's (1986) notion of "risk society" *{Risikogesellschaff).*
3. Weinberg (1972) has discussed the problems which arise when scientists go beyond the limits of science and operate in the field of "trans-science", that is, in the realm of scientifically unanswerable questions. According to Weinberg the decision modalities of trans-science differ in principle from those of normal scientific practice. We assume, however, that Weinberg's premise of an academic "republic of science" has been rendered invalid by history. Scientific learning takes place more and more by means of realistic experiments. For a critical view on Weinberg see also Jasanoff (1987).
4. Essential aspects of these reflections have been developed in an interpretation of the Chernobyl reactor accident by Krohn and Weingart (1987). See also the comprehensive study of Perrow (1984), who developed the notion of high-risk technology within the framework of organisational sociology.
5. Wynne (1988, page 158) tries to separate scientific from social experiments by pointing to the fact that "the 'hypotheses' being 'tested' (that is, those expert assumptions) are neither stated, nor even recognised. Therefore, by definition, they cannot be the subject of feedback and social learning". But one has to distinguish between assumptions implicit even for the experts (Wynne's focus) and those that are unknown to the public and only revealed by accidents. The cases discussed below reveal that, for the most part, a

well-defined research design exists, and learning by experimental implementation is a deliberate and scientifically performed research strategy.

6. On the differentiation between danger and risk, see Luhmann (1991).
7. Pasteur, 1922-1939, page 215.
8. The second alternative is the central theme of Latour (1983). MacKenzie (1987, page 214), too, has emphasised that technological systems tend to take control over their environments. See also MacKenzie (1989).
9. DFVLR, 1984, pages 74-75. Valk (1987) maintains from his experience in software/engineering that the reliability of complex technological systems such as space flight or SDI can only be tested by application.
10. Kaldor, 1981, page 134. She also points out that new aircraft are so complex that their design can never be complete (page 137). What she calls "baroque arsenal" refers to experimental implementations from a different perspective: since technological systems cannot be tested, customers for new weapons have no means of deciding about their practical value. They increasingly depend on the promises of the engineers who build ever more complicated baroque weapons systems.
11. See Bohme era/(1983).
12. MacKenzie, 1988, page 225.
13. On the basis of this insight Rip has attempted to create a new concept of TA under the designation "constructive technology assessment" (Rip and Belt, 1988; Rip, 1988; Rip, 1991). According to this concept (C)TA can be regarded as a "thematization process" which only empirically, that is by the actual construction of technology, may lead to the end aimed at.
14. The information in this paragraph is from the archives of the Union of Concerned Scientists; see Ford (1978, pages 10ff, 58ff).
15. Kunze and Eichhorn, 1973, page 336. For similar statements see the report of Kahlert (1988, pages 76ff).
16. See La Porte (1988).
17. See *Aviation Week and Space Technology* (AW&ST) (16 January 1989, pages 60-61; 30 May 1988, page 123; 6 June 1988, page 87); *Frankfurter Allgemeine Zeitung* (10-17 January 1989).
18. AW&ST, 16 January 1989, page 61,
19. In 1988 both engines of a Boeing 737-300 lost power and caused the plane to perform an unpowered landing (AW&ST, 30 May and 6 June 1988). This accident refuted the experts' assumption that a simultaneous failure of both engines was so unlikely as to be negligible.
20. *Frankfurter Allgemeine Zeitung,* 20 January 1989, page 15.
21. See Parnas (1986); Nelson and Redell (1986); Valk (1987); Un (1985).
22. See *Der Spiegel,* 11 July 1988, page 116.
23. Lin, 1985, page 39.
24. See *Frankfurter Allgemeine Zeitung* (26 March 1986, page 3).
25. AW&ST, 25 May 1987, page 25; compare with AW&ST, 2 June 1986, page 63.
26. AN/SLQ-32 has similar features to the more advanced *Aegis.*
27. AW&ST, 22 June 1987, page 32 (quote from the report).
28. AW&ST, 29 August 1988, pages 21-22.
29. See Perrow (1984, pages 133-134); Rochlin (1991).
30. AW&ST, 18 July 1988, page 23.
31. *Der Spiegel,* 11 July 1988, page 117.
32. *Frankfurter Allgemeine Zeitung,* 4 August 1988; AW&ST 29 August 1988, page 21.
33. See Weyer (1994).
34. Counter admiral Eugene Carroll (ret), according to *Der Spiegel* (11 July 1988, page 115).
35. The classical source is Carson (1962). A more recent study has been done by Dunlap (1981).
36. Dunlap, 1981, page 54.
37. See Trost (1984, page 64).
38. Subcommittee on Energy Conservation and Power (SECP), 1986, page 2.
39. SECP, 1986, page 23.
40. SECP, 1986, page 28.
41. SECP, 1986, page 34.
42. SECP, 1986, pages 2-4.
43. The main sources for the following are Trost (1984) and *Der Spiegel* (29 July 1991, pages 102-114 and 5 August 1991, pages 106-121), "Der Tod aus Ingelheim. Akte Boehringer: Wie Dioxin zur Waffe wurde".
44. Trost, 1984, pages 54-60.
45. Generally, the potentially hazardous materials used in the chemical and pharmaceutical industries should have to be classified in Schneiderman's (1980, pages 19ff) category of "uncertain risks".
46. Trost, 1984, pages 89ff.
47. Matthew Meselon, Harvard biologist, according to Trost (1984, pages 116ff).
48. Quoted from Trost (1984, page 158).
49. Trost, 1984, page 189.
50. Quoted from Trost (1984, page 302).
51. For an overview see Krohn, KLippers and Paslack (1987).
52. Stolarski, 1988, page 26.
53. See Turco et al (1984).
54. Trageser, 1989, page 38.
55. Trageser, 1989, page 39.

References

J Asdonk, U Bredeweg and U Kowol (1991), "Innovation als rekursiver Prozess. Zur Theorie und Empirie der Technikgenese am Beispiel der Produktionstechnik", *Zeitschrift für Soziologie,* 20, pages 290-304.

U Beck (1986), *Risikogesellschaft. Auf dem Weg in eine andere Moderne* (Suhrkamp, Frankfurt).

G Böhme *etal* (1983), *Finalization in Science. The Social Orientation of Scientific Progress* (Reidel, Dordrecht/Boston/ Lancaster).

R Carson (1962), *Silent Spring* (Mifflin, Boston).

DFVLR (1984) Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt, *Strategiestudie Raumfahrt* (Köln).

T R Dunlap (1981), *Scientists, Citizens, and Public Policy,* (Princeton University Press, Princeton).

D F Ford (1978), A *History of Federal Nuclear Safety Assessments: From WASH-740 Through the Reactor Safety Study* (The Union of Concerned Scientists).

R Herbold, W Krohn J Weyer (1991), "Technikentwicklung als soziales Experiment", *Forum Wissenschaft,* 8(4), pages 26-32.

R Herbold, W Krohn and J Weyer (1992), "Technisches Risiko und offene Planung. Strategien zur sozialen Bewältigung von Unsicherheit am Beispiel der Abfallbeseitigung", Bielefeld (unpublished paper).

S S Jasanoff (1987), "Contested boundaries in policy-relevant science, *Social Studies of Science,* 17, pages 195-230.

J Kahlert (1988), *Die Kernenergiepolitik in der DDR. ZurGeschichte uneingelöster Fortschrittshoffnungen* (Verlag Wissenschaft und Politik, Köln).

M Kaldor (1981), *Rüstungsbarock. Das Arsenal der Zerstörung und das Ende der militärischen Techno-Logik* (Rotbuch, Berlin).

W Krohn, G Küppers and R Paslack (1987), "Selbstorganisation - Zur Genese und Entwicklung einer wissenschaftlichen Revolution", in S J Schmidt (editor), *Der Diskurs des Radikalen Konstruktivismus* (Suhrkamp, Frankfurt) pages 441-465.

W Krohn and P Weingart (1987), "Nuclear power as a social experiment - European political 'fall out' from the Chernobyl meltdown", *Science, Technology and Human Values,* 12(2), pages 52-58.

B Kunze and H Eichhorn (1973), "Störfallanalysen von Kernkraftwerken", *Kernenergie,* 16, pages 333-336.

T R La Porte (1988), "The United States air traffic system: increasing reliability in the midst of rapid growth", in R Mayntz/ and T P Hughes (editors), *The Development of Large Technical Systems* (Campus, Frankfurt-am-Main) pages 215-244.

B Latour (1983), "Give me a laboratory and I will raise the world", in K D Knorr-Cetina and M Mulkay (editors), *Science Observed. Perspectives of the Social Studies of Science* (Sage, London/New Delhi/Beverly Hills) pages 141-170.

H Lin (1985), "The development of software for ballistic-missile defense", *Scientific American,* 253, December, pages 32-39.

N Luhmann (1991), *Soziologie des Risikos* (de Gruyter, Berlin).

D MacKenzie (1987), "Missile accuracy: a case study in the social process of technological change", in W E Bijker, T P Hughes and T J Pinch (editors), *The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology* (MIT Press, Cambridge/London) pages 195-222.

D MacKenzie (1988), "Ballistic missile stellar-inertial guidance: a study in the sociology of military technology", in E Mendelsohn, M R Smith and P Weingart (editors), *Science, Technology and the Military. Sociology of the Sciences, A Yearbook* (Vol XII) (Reidel, Dordrecht) pages 187-241.

D MacKenzie (1989), "From Kwajalein to Armaggeddon? Testing and the social construction of missile accuracy", in D Gooding, T Pinch and S Schaffer (editors), *The Uses of Experiment: Studies of Experiment in the Natural Sciences* (Cambridge University Press, Cambridge) pages 409-435.

T Maunz et al (1983), *Kommentar zum Grundgesetz* (Beck, Munich).

G Nelson and D Redell (1986), "Could we trust SDI- software?", in J Tirman (editor), *Empty Promise. The Growing Case against Star Wars* (The Union of Concerned Scientists, Boston) pages 87-106.

D Parnas (1986), "Software wars. Ein offener Brief", *Kursbuch,* 83, pages 50-69.

L Pasteur (1922-1939), *Oeuvre Completes,* vol VII (Paris).

C Perrow (1984), *Normal Accidents. Living with High-Risk Technologies* (Basic Books, New York). J

A Rip (1988), "The interest of the Netherlands Organisation for Technology Assessment (NOTA) in studies of science and technology", paper presented at the 4S/EASST Conference Amsterdam, 16-19 November 1988.

A Rip (1991), "The danger culture of industrial society", in R E Kapseron and P J M Stallen (editors), *Communicating Risks to the Public* (Kluwer, Dordrecht) pages 345-365.

A Rip and H vd Belt (1988),Consfructive *Technology Assessment: Toward a Theory* (Twente University, Amsterdam, manuscript).

G I Rochlin (1991), "Iran Air Flight 655 and the USS Vincennes: complex large scale military systems and the failure of control", in T R LaPorte (editor), *Social Responses to Large Technical Systems. Control or Anticipation* (Kluwer, Dordrecht/Boston/London) pages 99-125.

M A Schneiderman (1980), "The uncertain risks we run: hazardous materials", in R Schwing and W Albers (editors), *Societal Risk Assessment. How Safe is Safe Enough?" (Plenum Press, New York/London) pages 19-37.*

R S Stolarski (1988), "The Antarctic ozone hole", in *Scientific American,* 1, pages 20-26.

Subcommittee on Energy Conservation and Power of the Committee on Energy and Commerce, US House of Representatives (1986), *American Nuclear Guniea Pigs: Three Decades of Radiation Experiments on U.S. Citizens* (US Government Printing Office, Washington).

G Trageser (1989), "Idee des nuklearen Winters durch Waldbrande bestätigt", *Spektrum der Wissenschaft,* January, pages 38-39.

C Trost (1984), *Elements of Risks. The Chemical Industry and its Threat to America* (New York.

R Turco et al (1984), "The climatic effects of nuclear war", *Scientific American,* 2, pages 23-33.

R Valk (1987), "Schnell, genau und fehlbar. Warum Computersysteme nicht immer vertrauenswiirdig sind", *Süddeutsche Zeitung,* 17-18 October.

A M Weinberg (1972), "Science and trans-science", *Minerva,* 10, pages 209-222.

J Weyer (1994), "Actor networks and high-risk technologies: the case of the Gulf War," *Science and Public Policy,* 2(16), December 1994 (forthcoming).

B Wynne (1988), "Unruly technology: practical rules, impractical discourses and public understanding", *Social Studies of Science,* 18, pages 147-167.