TECHNOLOGY & THE FUTURE

Reader for WM0908TU

Dr.ir. Karel F. Mulder
Technische Universiteit Delft,
Faculteit Techniek Bestuur en Management,
Sectie Technology Assessment,
Jaffalaan 5
2628 BX Delft
tel 015-27 81043
e-mail: k.f.mulder@tbm.tudelft.nl

Table of contents

Why this subject?

Technology is a very strong force in determining our future fate. It has enabled us to live at an affluence level that is unique in the history of mankind.

But this achievement did not come without costs. Many new problems emerged by the advancement of new technologies, like environmental problems and privacy problems.

Companies are often depending on technological innovation to secure their survival. Nations face economic problems if they are unable to compete with the technological performance of their neighbors. Decisions on research, innovation and technology are crucial for national governments as well as (multinational) corporations.

To decide one has to be able to draw a picture of the future. What are the main social and technological trends? What will be the impact of these trends? Can they be steered? By who? Engineers that want to develop their careers beyond the engineering craftsmanship will have to deal with these questions. In this subject, we will deal with the question if we can find an answer to these questions and if so, which tools could we apply for that?

TECHNOLOGY AND THE FUTURE

Background

Following the industrial revolution of the late eighteenth and early nineteenth centuries, the next significant change in industrial production methods came around 1900. Until then, the innovation in products and production methods was chiefly the domain of individual inventors, whose technical ability was frequently accompanied by a marked commercial flair. The prime example is, of course, Thomas Edison. After 1900, however, the inventors' function was largely assumed by industrial 'in house' laboratories, such as those of the large chemicals companies (Bayer, BASF, Hoechst, DuPont, and later the Dutch AKU) and the electrical and electronics industry (General Electric, Westinghouse, Siemens and Philips)[1]. These laboratories applied the latest scientific insights and scientific methods. Some achieved notable successes: DuPont's invention of nylon and synthetic rubber, General Electric's development of the thermionic valve, and Bayer's discovery of Aspirin, for example. Between the two World Wars, the number of researchers on the payroll of private sector companies increased enormously, even after the Great Depression of 1929. In the USA alone, the number of industrial researchers rose from 8,000 in 1920 to 17,000 in 1927, and by 1938 had reached some 42,000.[2]

During the 1920s and 1930s, ever greater sums were invested in the development of new technologies. Scientific discoveries in diverse areas became increasingly important. It was no longer possible to arrive at a reliable evaluation of future developments based on the opinions of a single expert. 'Technology forecasting', an activity which had previously been undertaken *en passant* by managers and civil servants, became a discipline in its own right. Various technological options (each with varying chances of success) had to be weighed one against the other, whereupon Research & Development could then focus on the most promising. In many cases, the criteria applied were either further to military use or of commercial relevance. Technological breakthroughs such as television, for example, would open up a vast range of commercial opportunities.

Technology forecasting was first applied on a large scale in the United States. The resultant reports were often purely scientific in nature: they attempted to predict the course of a certain technological development in itself, with no regard for any social or economic implications. There have been frequent predictions of when the first manned space flight to the moon or Mars would become possible, when effective treatments for various major diseases will become available, about the car of the future, the house of the future, and so on. In most cases, such predictions were based on an (implicit) deterministic view of technology: a vision in which technological developments have their own dynamic and must therefore be seen in isolation from social or economic developments. In this vision, new technology has no social *causes*, but only social *effects*. Technology forecasts were therefore not greatly concerned (if at all) with such social effects. Their primary purpose was to set out the future in a way which would allow managers and policy-makers to determine their strategy accordingly.

In a few isolated cases, technological and social developments were indeed integrated within such a forecasting study. In 1944, for example, Theodore Von Karman produced a report

---

[1] Mulder, Karel F., 1992, *Choosing the corporate future. Technology networks and choice concerning the creation of high performance fibre technology* (Dissertation), Groningen.

[2] Landes, David S., 1969, *The unbound Prometheus, technological change and industrial development in Western Europe from 1750 to the present*, Cambridge: Cambridge University Press, p. 482.

entitled *Towards New Horizons,* in which he not only set out the technological possibilities of jet aircraft (based on the prior development of the gas turbine) but presented ideas for the way in which the US Air Force should function, politically and organizationally, after the Second World War.[3]

It was during the post-war years that the practice of technology forecasting really took off. Large institutes were created solely to consider what the future would bring.[4] In most cases, these ‘thinktanks’, the best known of which were RAND and HUDSON, took a particularly optimistic view, regarding the developments they foresaw as entirely unproblematic. The social desirability of a specific new technology was rarely considered. In most cases, it was tacitly assumed that the development of a new technology was not a 'zero-sum game' (i.e. that the any disadvantages would be in perfect balance with the advantages), but that society as a whole would definitely be better off for it. Of course, certain new technologies did indeed bring disadvantages, such as loss of employment or an erosion of power and influence. However, such disadvantages were accepted as the price of progress. Because new technology was therefore synonymous with social progress, it was considered unopposable. One could simply not object to progress!

During the 1960s, technology forecasts intended to predict future developments attracted much criticism. They were seen as evidence of the ‘technocratization’ of society. For example, new developments in housing (such as the ‘new town’ and the high-rise tower) which attempted to address the purely physical human requirements in an efficient way, but which also served to increase isolation and social disjuncture, were held in particular contempt. Opposition was based on the humanistic and social-democratic perspective. (A far left-wing stance was clearly not an alternative, since the left was also seen to favour rigid planning, as demonstrated by the enormous planning bureaus of the former communist countries).[5]

Nevertheless, the wider view of technology in general did change radically in the 1960s. Technology often showed itself to have unexpected and unwelcome effects, not only for the users of more outdated alternatives but also for society as a whole. New technology would sometimes bring advantages for those who developed it, but then later prove to have major disadvantages once it had been adopted on a wider scale. A prime example is the development of (nuclear) armaments. While the new weaponry undoubtedly increased the security of the country that developed it, the world as a whole became less safe. Similarly, new industrial production methods brought greater prosperity and employment, but often also eroded public health and well-being due to increased pollution[6]. It was therefore by no means certain that new technology would always and by definition have a positive social effect. The 'positive sum' game did not always apply. It became necessary to identify exactly who should pay the costs of the new technology and who should enjoy the benefits. But if the overall ‘costs’ were not in balance with the ‘returns’, what then?

The negative effects of technological development were often second or third order effects, i.e. those which had not been foreseen but resulted from a change in human behaviour further to the introduction of the technology.

---

[3] A.M.J. Kreykamp, H. Van Praag, B. van Steenbergen (redactie), 1972, *Toekomstonderzoek, theorie en praktijk* (1.2.2.), november, Deventer: Kluwer.

[4] E.g. Jungk, Robert 1954, *Tomorrow is already here. Scenes from a Man-made World*, London: Hart Davis.

[5] Waskow, quoted in A.M.J. Kreykamp, H.van Praag, B.van Steenbergen (eds.), 1974, *Toekomstonderzoek, theorie en praktijk* (supplement 10, 1.5.2.), November, Deventer: Kluwer.

[6] *Silent Spring* van R.L. Carson en L. Darling, (1962, Riverside Press, Cambridge, Mass.) was particularly influential.

> The first order effect of the introduction of 'the pill' was a more reliable, more convenient and more pleasant method of contraception compared to the condom. A second order effect was greater sexual freedom. A third order effect was then an increase in the incidence of venereal disease, which eventually prompted many people to resume using the condom (since greatly improved): a fourth order effect.

Based on these insights, it was no longer clear how new technology should be evaluated. To what extent could the social and societal effects be predicted? It also became clear that technological developments themselves could be strongly influenced by social circumstances. (An example would be the extensive energy efficiency research programmes prompted by the oil crisis.) If the effects of technology could be foreseen, how could the development process itself be influenced and directed? A new type of study was introduced to answer this question: the *Technology Assessment* (TA). However, the exact form and content of a technology assessment remained unclear for some time.

> ***"Technology assessment is an attempt to establish an early warning system to detect, control, and direct technological changes and developments so as to maximize the public good while minimizing the public risks."***[7]

Under this definition, the TA was primarily an instrument to meet the needs of the (critical) public. However, other definitions left the decision-makers themselves to address the political pitfalls presented by the use of terms such as 'public good' and 'public risk'. The definition offered by Emilio Daddario, who was responsible for the bill under which the US Congress Office of Technology Assessment was established, regards TA specifically as a policy instrument:

> ***Technology Assessment is a form of policy research which provides a balanced appraisal to the policymaker. Ideally, it is a system to ask the right questions and obtain correct and timely answers. It identifies policy issues, assesses the impact of alternative courses of action and presents findings. It is a method of analysis that systematically appraises the nature, significance, status, and merit of a technological program.***[8]

Although these definitions differ greatly in the way in which they demarcate the role of TA within political decision-making, they also have a clear common denominator in terms of the way in which they set TA apart from previous technology forecasting processes. There are two main differences:

- TA is not primarily interested in the development of technology itself, but in the development of technology in relation to the social context, i.e. the manner in which social choices are made (and will be made), and the acceptance and appreciation of the possibilities of the new technology on the part of various societal groups.
- TA is, to a greater or lesser extent, concerned with identifying the opportunities for influencing the course of technological development, both in the positive and negative sense, and hence not with any accurate prediction of the technological future. Whether TA should itself adopt some standard ('minimizing public risks') or should

[7] Marvin J. Cetron, Lawrence W.Connor, 1972, "A method for planning and assessing technology against relevant national goals in developing countries", in: Marvin J. Cetron, Bodo Bartocha, *The Methodology of Technology Assessment*, New York.

[8] E.Q. Daddario, in *Subcommittee on Science, Research and Development of the Committee on Science and Astronautics, US House of Representatives, 90th Congress, 1st Session, Ser. 1*, (Washington DC, US Government Printing Office, Revised August 1968), p. 10.

leave this to the policy-makers is subject to discussion.

Technological forecasting in general has certainly been influenced by the emergence of TA and by criticism of the traditional approaches. Today, technological forecasts are often made within the framework of the somewhat broader terms of reference of TA. Accordingly, we shall focus on TA itself, although this should not be taken to imply that technological forecasts are only relevant to TA.

Types of Technology Assessment

A number of different types of TA studies can be identified, based on their phasing in relation to the decision-making process:

ATA, Awareness TA: an assessment of possibilities. This can, in principle, be conducted regardless of current political relationships, since the aspects to be studied do not have to be determined in advance. Forecasting methods such as Delphi play an important role.

STA, Strategic TA, is intended to inform a specific actor about his strategic options with regard to technological potential. However, the analysis is not confined to that actor's strategy, but considers the sector as a whole. The STA comprises an analysis of:

- threats and opportunities
- the objectives to which the technology may be able to contribute
- the relevant preconditions
- barriers within the system
- socio-institutional innovations.

CTA, Constructive TA: focuses on the improvement of the process of technological innovation. Social relationships are developed, the opportunities for technological change are set out. It is not intended to develop socially 'useful' products, but to identify the opportunities for the development and strengthening of relationships between various stakeholders.
It will:

- contribute to the creation of an infrastructure
- include experiments and trial projects to promote interaction between R&D and actual application
- collect information and experiences from the experiments in order to further learning processes.

It is subject to certain limitations in that:

- unpredictable effects can become dominant
- it may not be possible to reach consensus on all aspects. Where no comparison for certain hazards can be found, for example, there is no modality for CTA.
- Not all interaction processes will automatically involve dialogue, whereupon people may work at cross purposes.

- Not every desirable technology will attract funding (disparity between social demand and affordability).

Backcasting is a form of TA study with an extremely specific objective. First, a target situation which should be achieved by a certain date is defined. The analysis then considers how this situation can be attained: what relationships must be developed, what technological developments are desirable, and which are not? The degree to which the defined situation differs from the existing situation determines the number of opportunities to bring about the desired change.

TA institutes

Government organizations which can act in a directorial capacity are themselves also technology *developers*. The strategic information about the effects of technological developments has therefore also played a role in territorial disputes between the legislative bodies (parliaments) and the implementing authorities (governments). Parliaments have traditionally regarded information about the likely social effects of technological developments as extremely pertinent, and such information can prompt them to introduce timely legislation. For the implementing authorities, this is less desirable, since developments which would normally not fall within the domain of parliamentary decision-making could now be subject to political debate. Accordingly, it was more or less inevitable that the parliament of the most technologically advanced country would be the first to establish an official TA organization: the US Office of Technology Assessment, which soon became an official agency of Congress[9]. The founding of the OTA is inextricably linked with the name of Senator Emilio Q. Daddario, who not only succeeded in having the necessary bill passed, but became its first director in 1972. His tenure was not a conspicuous success, since he was inclined to avoid any controversial issues in order to placate Congress, on which the OTA's very existence depended[10]. Later, however, the OTA (under a different director) managed to develop an extremely good reputation. Even so, it fell victim to a massive cost reduction programme and was disbanded in 1995.
In most European countries (and in the EU itself) Technology Assessment only really established itself in the 1980s. There had been some initiatives during the preceding decade, but these failed due to the European parliamentary tradition (in which there is no distinction between legislative and implementing authorities) and the perception of TA as a strengthening of parliament's position at the expense of government, and hence a strengthening of opposition.[11] The Netherlands Organization for Technology Aspects Research (NOTA) was founded in 1986[12], later changing its name to the Rathenau Institute in 1994. At around the same time, several other technology assessment organizations were

---

[9] Gray, Lewis, 1982, "On 'Complete' OTA Reports", *Technological Forecasting and Social Change* 22, 3 and 4, pp. 299-319.

[10] Dickson, D., 1984, *The new politics of science*, Pantheon New York.
For an account of the OTA's methods, see Fred B.Wood, 1982, "The Status of Technology Assessment, A view from the congressional Office of technology assessment", *Technological Forecasting and Social Change* 22, 3 & 4, pp. 211-222.

[11] Smits, Ruud and Jos Leyten, 1991, *Technology Assessment, waakhond of speurhond? Naar een integraal technologiebeleid*, Zeist, Kerckebosch.

[12] Smits, Ruud, Arie Rip, 1988, "De opkomst van TA in Nederland", *Wetenschap & Samenleving* 40, no. 5, pp. 7-16.
Maarten Evenblij, 1989, "Technology Assessment", *Intermediair* 25, nr. 7, 17 February, pp. 33-37.

created throughout Europe. They included FAST (EG), STOA (European Parliament), OPECST (France), TAU (Austria), SERV (Flanders), TAB (Germany), POST (United Kingdom), Technologi Naevnet (Denmark), and IFS (Sweden).[13]

TA and technology forecasting in the private sector

TA has also established a position within the private sector, due to several reasons. Firstly, the traditional technology forecasts often proved inadequate. It gradually became apparent that not everything which can be described as 'technological progress' will also be commercially viable and profitable. For example, Concorde proved that supersonic air travel was indeed technologically possible. However, the emergence of wide-body aircraft made low-price air travel viable. Rising oil prices hit the kerosene-guzzling Concorde particularly hard. Moreover, it was difficult to obtain landing rights for Concorde due to its negative environmental impact. As a result, Concorde was a commercial failure. In fact, all these factors could have been foreseen, rendering this one of the clearest examples of a case in which technological forecasting studies should have been broadened to include a full consideration of social and economic factors.[14]

In the private sector, TA is primarily concerned with the social effects of a company's products and the opportunities for the development of new products to address a certain demand, or indeed to create that demand. It has also been applied to predict government responses, scarcity situations, etc. This form of TA is usually conducted 'behind closed doors', since it is not concerned with any broadening of the decision-making base by sharing information, but solely with the attainment of competitive advantage by being 'the first to know'.[15]

TA and forecasting

The claims implied by the various definitions of TA have been proven largely unfounded in practice. Gradually, serious doubt emerged as to whether it really is possible to create an effective **early warning system** for technological developments. Moreover, to predict all social effects of technological development (including the third and fourth order effects) proved almost impossible, particularly given the rapidly changing social conditions. For example, the oil crisis coincided with, and thwarted, many TA analyses. In response, the objectives of TA studies often had to be defined far more narrowly.

This prompted the definition of a new form of TA, which differs from the traditional TA in the following respects:[16]

TRADITIONAL	**NEW**
Dominant role for science	Researchers and users are equal
High expectations with regard to the possibilities of TA	Modest expectations with regard to study results

[13] Smits, R.E.H.M. 1990, *State of the art of Technology Assessment in Europe*, A report to the 2nd European Congress on Technology Assessment Milan, 14-16 November 1990.

[14] Jones, Harry, Brian C. Twiss, 1978, *Forecasting technology for planning decisions*, The Macmillan Press Ltd., London/Basingstoke.

[15] Simonse, A.S, W. Kerkhoff, A. Rip (eds), 1989, *Technology assessment in ondernemingen*, Deventer, Kluwer.
Wissema, J.G., 1977, *Technology Assessment: aspectenonderzoek in het spanningsveld van technologie en samenleving*, Deventer, Kluwer.

[16] Smits R., J. Leyten, 1988, "Key issues in the institutionalization of TA", *Futures*, February.

Output = report	Output = report and discussion
Problem definition of secondary importance	Consideration attention for problem definition
A single TA organization	Many TA organizations
Instrumental use of information in a rational decision-making process	Conceptual use of information in political decision-making process
TA results automatically included in decision-making	Considerable attention devoted to matching TA to the decision-making process
Technology is autonomous	Technology is made by people

The debate surrounding TA has also had a profound influence on the practice of technological forecasting. On the one hand, the methods and techniques of forecasting often play an important role in TA studies. On the other, technological forecasts now often include considerable attention for social elements. However, the main difference between the two types of study remains the *purpose*. Do we wish to be able to respond promptly and adequately to future technological advances, which often cannot be influenced (at least by us), or are we to try to render technology fully capable of social influence? It should be clear that these questions are so closely interrelated that they cannot be answered separately. The Technology and the Future course module (WM0908) therefore focuses on the possibilities (and impossibilities) of forecasting future technological developments, while Technology Assessment (WM0909) looks at the possibilities for directing and influencing technology.

TECHNOLOGICAL FORECASTING: 1970-1993', Joseph F. Coates, John B. Mahaffie, and Andy Hines[17] Technological Forecasting and Social Change, No. 7, pp. 23-33 (1994)

Introduction

Under the sponsorship of 18 large organizations, Coates & Jarratt, Inc., conducted Project 2025, looking to how science and technology will affect the United States and the rest of the world over the next generation. In the first phase of that project, we collected all the science and technology forecasts we could find done since 1970 and projecting any time forward from that year. The search was organized in 54 scientific and technological areas in order to cover forecasting in all of science and technology. In a second phase not reported on here, we created our own forecasts of the year 2000.

The results were presented in 41 reports, each of which defined the principal anticipated outcomes and the capabilities, which were anticipated to be delivered to society. We identified gaps and points overlooked in the forecasts. This was easier to do with regard to the earlier forecasts. We also identified the scientific, technical, and social assumptions underlying the forecasts. It came as no surprise that little by way of assumptions was explicit in the forecasts. We often had to impute assumptions that must have been made in order to come to the forecasted conclusions. Our work then proceeded to identify both business and public policy implications in each of the reports. Finally, we presented a digest of the main forecasts and our pithy evaluations of them.

The undertaking was global in scope, that is, not limiting us to the United States or English language forecasts. It quickly became clear that we could not accomplish our task by limiting the search to formal forecasts. We had to expand the search to include two surrogates for forecasts. One was research agendas. This is based on our assumption that if there is a research agenda, that research will get more attention than other topics and hence will lead to more practical or applied results. The second surrogate was critical technology agendas, again operating on the assumption that what is identified as critical is likely to get more attention than other subjects in the same field.

We also augmented formal forecasts with a large number of more or less incidental forecasts often made in connection with a speech, a journal article, or a semi-popular publication.

The State of Forecasting

Overall, we see forecasting as underdeveloped. It was better developed in the 1960s and has decayed in methodological quality and substantive content. The more recent forecasts are more often informal, side commentaries, or poorly defined and executed without much attention to assumptions, time horizons, or the author's intentions.

On net, all too often the forecasts we examined did not give their rationale and did not explain their assumptions well. They often did not identify the time horizon at all. They were not good about explaining the capabilities at the core of the technological development, i.e., at defining what the technology would allow us to do.

In setting out to do this project, we thought that it would be a daunting task. It was. But daunting for the wrong reasons. There were far fewer forecasts than we expected to find. That was not an artifact of our approach. There had not been the flowering of forecasting in the

[17] Joseph F. Coates is President of Coates & Jarratt. Inc. John B. Mahaffie is an Associate with Coates & Jarret, Inc. Andy Hines is an Associate with Coates & Jarratt, Inc. Address reprint requests to Joseph F. Coates, Coates & Jarrett Inc., 3738 Kanawha Street, NW, Washington, DC 20015. (website: http://www.coatesandjarratt.com/)

1970s and 1980s we thought there had been. The quality of forecasting is very, very mixed. There are fields with next to no forecasts and others with rich, regular, frequent, formal, rigorous, quantitative forecasts. In aerospace and information technology, there is widespread industrial and governmental emphasis on forecasts. They do a great job at it. In other fields, such as economics and basic mathematics, there is little or nothing.

Sometimes people outside a field have more interesting things to say about the field than the insiders do. Sometimes the visionaries are not in the center of the field but people who look across all science and technology and think broadly and liberally about what could be. Those visionary forecasts are not necessarily always rigorous or quantitative, but often are more interesting and useful than institutionally-based forecasts.

We often found less formal forecasting interesting where people comment on the direction of technology in the context of some other thing that they are doing. A common example is people who are looking at the future of a profession from the point of view of the supply and demand for the professionals. They sometimes turn to thinking about where their science is headed. An undeveloped aspect of forecasting is putting expectations about a profession together with the forces inside and outside the profession that are shaping it.

Some fields have done more with forecasting the future of the profession than they have with the future of the science and technology. An example is architecture, where there are many forecasts about the fees, business opportunities, and ways of keeping the architect's grip on the action in the face of changes like CAD/CAM.

Technology forecasts that are very specific about some aspect of a technology, e.g., the number of transistors on a chip, are common. Less common are broadbased looks at a whole field, its related fields, and the social contexts surrounding them. These minutia forecasts are often at the expense of a careful look at what might completely upset the whole field. The concentration in forecasts is most often on marginal or incremental changes.

In close-knit fields forecasts often show a great deal of consensus. They forecast the same thing down through the years. So there is some danger that a tightly-knit field misses the broader possibilities because they only read each other's work. Forecasts in food science are an example of this. If one changed the dates and a small amount of language of many forecasts made 10 or 15 years ago, they would have a striking resemblance to more recent forecasts. On the other hand, this may not reflect exceptional narrowness coming from merely talking to one's self. It may also reflect the slow and steady pace of a large sector of the technologically-based economy.

Regrettably, there are a lot of things posing as forecasts that are not forecasts. Numerous journal articles have a catch phrase, such as "past, present, future" or "yesterday, today, and tomorrow." All too often they deal with past and present but offer little or nothing about the future. We conjecture that these misleading articles are often done by specialists in the field who implicitly have a model of continuity, a model of technological momentum in which continuity dominates over change. Furthermore, many of those articles are written by people who had little opportunity to formally explore a futures paradigm and, hence, just do not know how to approach the forces and factors shaping the future of the subject of their concern. It surely suggests some interesting opportunities for professional societies in almost every field to educate their members on how to think about the future. Technology enthusiasts and visionaries often see their technology as the one that will be the hottest new thing in the years ahead.

In contrast to the Japanese who make a crucial business point of normative forecasts in setting goals and direction, the publication, celebration, and policy use of normative forecasts in the United States is extremely limited. The Department of Defense, as well as NASA, has for years used normative forecasting to help shape the next wave of technological development. Aside from that, relatively little forecasting occurs as a conscious social steering mechanism in the United States.

Forecasting too often mixes technological and market forecasting. People who have a particular product to market may steer themselves down the wrong path because of their overwhelming interest in the market. They in essence let market expectations drive expectations for the technology.

In reviewing the 54 areas in which we gathered forecasts, four clearly stood out as the best: aerospace, information technology, manufacturing, and robotics. Similarly, six areas were conspicuous for the paucity of forecasts and their general poor quality. Economics, as well as most of the rest of the social sciences, was very weak, as were physics and basic mathematics. It is interesting to note, in sharp contrast to basic mathematics, people in statistics have a good history and pattern of forecasting. Zoology and botany, that is, general biology, were weak in contrast to modern molecular biology and genetics. Finally, geology and soil science and related areas were also relatively uninteresting.

Why Is Forecasting So Uneven?

While it would be difficult to be definitive about the reasons for these clear patterns, there are some suggestions. First, when there is a technologically oriented sponsor who has a strong economic interest in the subject, there tends to be a good bit of forecasting. This would surely characterize the four leading areas we have noted. On the other hand, when the issue is politically charged or when no one has a particularly strong economic interest in forecasting, there are few forecasts.

However, there are ironies. Economists who forecast all the time have fairly consistently avoided forecasting about their own field. Basic mathematics and physics are extremely esoteric fields pursued by a relatively small coterie of extremely intelligent people. We suspect that there is some arrogance as well as a degree of intellectual caution that retards forecasting for them. There seem to be no obvious reasons for the dearth of solid forecasting in the social sciences. They do, however, seem to be increasingly driven by ideological and political concerns, as well as a hefty move toward social action agendas. These trends discourage forecasting. The four other sciences weak in forecasting comprise the routine core or background to applied areas and, hence, have no particularly strong clientele. For years, we have tried to get the US Geological Survey interested in a forecast of the geological sciences and have consistently come up zero on that. Perhaps it is a case of what difference would it make? As the nation moves more and more toward an agenda of competitiveness, as government becomes more and more concerned about supporting the obviously central role of science and technology in our future prosperity, it is ironic that there is no clear government agenda and virtually no agency champion of a systematic approach to forecasting. This is in striking contrast to the situation in Japan.

Applied science fields forecast things more often and probably with better results than pure science fields. That is a shame because basic scientists need to know where they may be going in a world of increasing cross-disciplinary work. They know rather well what is happening around their immediate interests, but they do not necessarily think about what will happen in five, ten, or twenty years to reshape their field or the consequences of what will result from their research.

We have not been able to figure out why the basic sciences seem to be so resistant to forecasting. Possibilities include a fear of tipping one's hand, that is, revealing one's own research agenda, or maybe a fear that legislators or other sources of funding may find the anticipations uncongenial. Or, it may be an ironic anti-intellectual arrogance that leads basic researchers to believe that their fields cannot be forecast. A striking example of that is the difference between applied mathematics and pure mathematics. We found nearly no forecasts in pure mathematics. In a couple of interviews to search out forecasts, we got the foolish response that we did not understand that basic mathematics is so creative that one simply could not forecast it. On the other hand, the applied mathematicians, particularly in statistics, have a good record of forecasting.

The Four Enabling Technologies and an Enabling Issue

Four enabling technologies turned up over and over again in the forecasts in many fields. First is the broad family of information technologies. For obvious reasons, computers, computer networking, data, data gathering, telecommunications, and sensing are influencing every field. Forecasts in most fields saw information technology as shaping their fields.

Second is genetics and related biotechnologies which are increasingly prominent in forecasting. We saw a changeover in the late 1980s with more forecasts finding genetics relevant to their field. While not every field identifies genetics as relevant, a majority do.

Third is materials science and technology, which is critical to any field that manipulates things. Most of those fields recognize an emerging revolution in the materials entering into all structures, devices, and artifacts.

Fourth is energy technology. Behind a lot of forecasts, in many areas, is the expectation that we will have the energy we need in the form we need it at the price we can sustain for that technology. While many people see a need for radical transformation in the energy base of the United States and the global economy, there is surprisingly little by way of radical forecasting. Equally surprising is the relatively little systematic, comprehensive, in-depth, normative, i.e., goal-directed forecasting of the energy future. On the other hand, there is a great deal of emphasis on the forces and factors leading to new energy arrangements.

A fifth area, not so much an enabling technology but an enabling issue, is environmentalism. Nearly every science and technology field at some point recognizes the environment is critical to its future. There may be some that have not woken to that yell but they will.

By an enabling technology or an enabling issue, we mean one which has effects not only in the area to which it is immediately directed but one which brings about basic changes in many other areas. The electric light turning night into day had radical effects on the way we use time and space. Similarly, the automobile did not just substitute for the horse and buggy but spawned effects that created 10% of the national economy.

Problems in Forecasting

A curious finding all too common is that experts in a field often do not know about the forecasting in their field. For example, in a typical field, to find forecasts, we contacted between 15 and 40 people. At the same time, we did electronic and library searches on the future of the field. Between the experts and the fields' databases, we had the best possible coverage. It is surprising how quickly experts in a field can forget what somebody wrote about the future five, much less ten or fifteen, years earlier.

There definitely is a database search and nomenclature problem in identifying forecasts and futures research across the whole scientific enterprise. Bibliographers should be paying much more acute attention to the subject of the future. Often, forecast and future are not even used as key words and descriptions in the coding of literature for electronic databases.

Characteristics of the Forecasts

In our review of over 1500 forecasts we did not attempt to evaluate their reliability, that is, to what extent what they forecast and occurred. That, in itself, would have been another major project. However there are grossly visible patterns about the reliability and effectiveness of the forecasts. Perhaps the most common characteristic of forecasts in science and technology is incremental change. Incremental change did pretty well because within many fields, people know their business and their technology, and they know the possibilities a few years or a decade out. So forecasting from within a field taking into account what is going on in the field is probably the most successful form of forecasting in the relatively short-term, except where something external comes along to upset the apple cart.

An interesting example of combining continuity and change are the forecasts in microelectronics. They are numerous, frequent, and highly quantitative, and yet as new scientific and new technological developments occur as they have over the last 20 years, they have been effectively integrated into the forecasts. The forecasts in the field of microelectronics tend to reflect steady, rolling change, and because of the large number of forecasts and the continuity of forecasting, the evolution in anticipations shows up distinctly. On the other hand, in fields in which forecasts are incidental or spotty, the discontinuities, in the form of new developments, do not show up clearly or get effectively integrated into forecasts.

There seems to be an implicit view in many of the forecasts we looked at that forecasts are attempting to give a right answer or to correctly describe some future situation. Certainly among most futurists, this is not the intention of looking to the future. Rather, futurists try to define a range of alternative futures and to use that full range of alternatives as the basis for planning. We found very little laying out of alternative developments in the forecasts that we reviewed. Surely the most important measure of a good forecast is not whether it is right or wrong, but whether it pushes developments in a useful direction. Because we chose to avoid all proprietary forecasts, we may be missing a lode of highly successful forecasts. Unfortunately, we have no way of evaluating that situation. Apparently, most organizations choose not to encourage publication, even after years or decades, of their forecasts.

In some areas there is a kind of long-range optimism which never seems to be fulfilled so that the forecasts of 25, 20, 15, or 5 years ago all look the same. A most interesting example is that of fusion energy, which for the last quarter century has been always 50 years in the future.

Forecasting Around the World

Searching the world outside the United States for forecasts was disappointing. The futures community was not forthcoming. For example, we wrote to close to 200 fellow members of the World Future Studies Federation, which is the most broadly based futures society in the world. Its membership is carefully selfselected so that everyone is a legitimate futurist. From that total inquiry we got 3 responses, one of which was interesting and useful and another that was a plea for money. This weak response reflects a melancholy situation with regard to the systematic study of the future, namely that people are reluctant, even unused to cooperation. They do not approach their work on a professional basis with a sense of professional exchange. The lack of response to our project was equivalent to "I couldn't care less." Incidentally, one of us is a member of the Federation, so we were not approaching the other members as an outsider.

Most of the forecasting done regularly and in some detail is in the U.S., Japan, and Europe.

We did not get anything from the Third World. To some extent, Eastern Europe and China and Russia also do forecasting.
With regard to the EC activities, there is a complex, interlocked, cascading collection of materials. The European Commission has a number of offices that are charged with forecasting technology and with tracking forecasts in technology. One of us (J.M.) visited there and beat on doors all up and down the hallways. They have not gotten very far with science and technology forecasting. They have nice names for their various institutes and commissions, but they have quite a way to go. They have the resources and the people-power. They just have to get on with it. We do not believe that we were in any way excluded or denied material. Rather, what we think is that the system is not yet effectively organized to produce reliable, high quality forecasting products that the EC community has every reason to expect if not demand.
The OECD has resumed its considerable interest in the exploration of the future. Its work, however, is relatively new and started up too late to provide a significant input into our 2025 project. The OECD Future Studies Information Base is putting out occasional papers under the title HIGHLIGHTS with such subjects as world population, water, and other topics of general interest to the OECD. These are outstanding interpretive summaries of current literature, including forecasts and futures analysis with regard to each topic.
The Japanese forecasts are, without question, the most comprehensive, systematic, long-range, and sophisticated. Their commitment to forecasting began about 1975, and they have made particularly effective use of broad-scale, well done Delphi surveys. They enjoy a great deal of continuity and overlap from one study to another, and the studies are sponsored by organizations which are prepared to think about them and act upon them. The futures work has also begun to systematically permeate the Japanese professional literature. While a large amount of material is available in English, there was for us a substantial language barrier, so that we did not extract as much of the gold from the mine as we could have, had we the language capabilities for more translation.

Limitations of Observations

The Project 2025 material dealt exclusively with non-proprietary information, hence we excluded any discussion of classified government material, and we purposely avoided access to any internal corporate documents. Therefore, the extent of hard-core professionally excellent work may be greater than the above material suggests.

The Future as Pursued by Corporate and Government America

Technological forecasting is only one, albeit major, portion of the futures enterprise. There is therefore some value in looking at the larger pattern of trends in futures research in corporate and government America, to appreciate the shifting patterns of priorities and their wider embrace of a futures paradigm.
In several of our projects, including Project 2025, we have asked our clients what their experience had been with futures research and forecasting. To an overwhelming degree, we have found that they have been extremely dissatisfied with forecasts done 10 to 20 years ago. There are two separate but related reasons.
First, the users were often left unaware that there were fundamental scientific or technological assumptions made, which were unstated and hence unexamined, which turned out to be unstable. Second, and perhaps of even greater importance, the assumptions about

the state of the society - the corporate external environment into which the new development would be delivered - was itself often unexamined. One incidental consequence of that is, in all our work, we have been made aware of the need to make contextual assumptions as explicit as possible.

The above are likely reasons why technological forecasting and a general interest in future studies declined in business in the late 1960s, through the 1970s. We attribute the revival of interest in the future coming from two separate factors. First is that the corporation, whether American or foreign, is now caught up in an unprecedented degree of competitiveness. As a result, there is a widespread interest in virtually any technique or approach which promises to give insight into that competitive environment. The interest in the future is one of several areas that are prospering as a result of that concern. Separate, and distinct from that, is what we have come to call the "magic of the millennium." As the new millennium approaches, many people and organizations are behaving as if they feel that we are at an objective branch point, that at the millennium we will know whether America is on the right road or whether a particular corporation will succeed or fail.

Accompanying the revived interest in the future in the 1980s and early 1990s is a broad commitment to the communication of results. That is partly recognition of the need to tell the story well and partly a way to achieve more effectiveness in futures research through a greater commitment to client involvement with the study itself. The day is past in which a study will be completed, presented, and that is it. The best of work is done with extensive interaction with the client and with relevant parties at interest to assure maximum utility.

There is also a gratifying increase in the time horizon of futures research. In the early 1980s it was difficult to get anyone in business or government interested in more than three or four years because of the tremendous pressures for short-term return on investment. This was reflected in the foreshortening of the time horizon of much of corporate planning. We find more recently, as reflected in our Project 2025 and a current project looking at American business out to 2020, that it is no longer impractical to find active interest in the 24 and 30 year future.

There also is a general awareness in large organizations, both public and private, that the study of the future does have something to tell them. Accompanying this general awareness of the potential value of looking to the future is a melding or blurring of technological forecasting with a more general and often less quantitative look at forces and factors shaping any particular field of interest.

At the corporate level, we find the interest in the future not particularly high in strategic planning units, but we are finding interest in R&D, in advanced market research, and in human resources. We also find growing interest in the exploration of the future among the best of companies, which fear that they may have been talking to themselves too much to the exclusion of messages from the outside. There is a growing interest in outside inputs into their planning and strategic thinking.

This broad, diffuse interest in the future is nicely illustrated by a quotation from one of our clients in the utility industry:

> Knowing our customers has always been important. Now it's becoming absolutely crucial for us to understand their wants and needs. Tracking and studying established trends helps us think through the real needs and preferences of today's customers and anticipate future changes in customer attitudes and perceptions.

A human resource executive in one of the Baby Bells reports:

> Forecasts and futures research have proved to be the stimuli needed to get us, as an organization, to look beyond our own view of the world. All too often, we see our world with the bias of our problems, our industry. Forcing us to look beyond that bias causes us to challenge our assumptions and ourselves. There are few answers out there; however, there are tools that help get you closer. That is the role for forecasts and futures research.

A project manager in a manufacturing association finds:

> The use of futures forecasting is a key methodology for identifying long-term strategic thrust areas, which in many cases may be direct threats to existing businesses. As such, they provide direction for, and a sense of urgency to, longer-term research and manufacturing efforts. They can also help shape the types and backgrounds of people an organization hires over time to help lead it into new paradigms.

By no means, however, are these good feelings about the use of futures research universal. They vary not only by company within business sectors, but they vary by business sector themselves. The unpleasant reality is that some business sectors are tuned out of the need to look to the future.

A senior analyst in an energy corporation reports the following:

> With regard to the energy industry and forecasts, the tendency is very strong to look at the short-term forecasts of price and demand. There is, however, growing dissatisfaction that the users of those forecasts are not getting what they are buying. Futures research, except for E&P, in the energy game is unfamiliar. Essentially it is an a technological business, and so technological changes always come as a surprise. It is basically only those related to geology and more recently those connected with environmentalism who are beginning to look at the future. In summary, the industry just does not understand futures research.

A different realistic look at the use of futures work, given by a senior executive in a chemical company:

> The largest potential to use a long range futuristic forecasting of science and technology in the industrial community comes from a technology oriented company that is committed to growth by finding and developing business opportunities for new and advanced products or services. Project reports are useful in brainstorming and planning activities to select opportunity areas for a limited amount of long range corporate R&D.

Another factor which probably applies more specifically these days to the chemical industry than to others is the effect of environmental issues. An increasing share of capital investments and R&D budgets are used to address environmentally-related issues which leaves fewer financial resources to support research for other new product and processes. But the latter are a key to repositioning companies into new business areas that are being spawned by these forces. To this end, Project 2025 has offered valuable exposure to issues and opportunities in some fields that were relatively unfamiliar to us, and the potential to continue to use it this way remains.

The box summarizes the applications of futures research in one important component of a chemical company, Dow Ventures. The material is a direct quote from our client at Dow, Kerry Kelly. The material illustrates the importance of communication, the need for extensive and continuous client or user involvement with the work, and the problems and opportunities associated with broad dissemination of futures material in a very large organization. It also illustrates the requirement that information with implications for change come from multiple sources if it is to be credible.

Futures in Government

The story of futures in government is complex and checkered. Ironically, the Reagan administration, with its very unequivocal and strong anti-bureaucratic sentiments, was a strong stimulus to futures research in the federal government. The administration's position was that the bureaucrats should behave more like the big boys in business. When the bureaucrats looked around they found that one of the things the big boys did was strategic planning and futures research. More recently, with the approach of the new millennium and with the vigorous activity directed at reinventing government, almost all agencies have developed some kind of year 2000 initiative. Unfortunately, as near as we are able to tell, most of them are winging it, that is conducting their studies as internal activities with their own staff largely free of professional input from the futures research community. However, the story is mixed. Many agencies are using professional futurists, and some agencies have fully qualified and competent futurists on their staffs. The overall effect is that government at the federal level is steadily moving toward a greater awareness of the value of the systematic exploration of the future. The FBI and the EPA have done, or are engaged in, futures studies and programs.

The quasi-governmental bodies present a mixed picture. At the time of this writing, the Smithsonian Institution has a Commission on the Future of the Smithsonian made up largely of people in or peripheral to the museum field. On the other hand, the National Academies have been adamantly resistant, with few exceptions, to a serious and systematic embrace of the future. This is ironic since almost everything that the academies touch are important not because of the past or the present, but because of their implications for the future. For honorific organizations, a firm grasp of the future can be threatening.

Using Futures Research at Dow Ventures

We find most forecasts to be vague and supported by specific examples which may or may not be indicative of trends, rather than projections from statistical data which integrate many examples. As a result, it is difficult to build credibility in the organization for futures work. Consequently, the information is not integrated into the planning process. One of the best works we have seen in the future studies was Project 2025, which did a thorough map of several technologies and integrated them into a few functional scenarios.

We put copies of the reports into our Business Information Center with appropriate key words so anyone doing a literature search would find the appropriate reports. When we received the assumptions for phase II of the project, we began an e-mail survey by sending a few (2-3) of the assumptions at a time to over 100 R&D and Ventures personnel. The purpose of the survey was to test the believability of the assumptions and to begin to distribute the information more broadly and begin to get the organization thinking in future terms. This was very successful. The response rate was high and informal feedback *indicated that people were integrating the ideas into their thought processes which then became integrated into the business strategies and R&D programs.*

When the phase II reports were completed, we distributed them to the business teams and Ventures groups which were most directly aligned with the reports. They were asked to distribute them within their groups. In some cases, we had the authors review the reports with top Ventures management and lead a brainstorming session with a cross-functional and cross-business group to generate new business ideas. In some cases, phase II reports and some phase I reports were used as prework for brainstorming sessions conducted by our Chemicals New Business Development Group.

All of the Project 2025 materials are in our Business Information Center in Midland, Ventures, Chemicals & Performance Products New Business Development, Plastics New Business Development, and Dow Europe. These reports are used as reference materials when we begin work in new areas and as an input into business strategies for our new business development activities.

We will also be using future studies to identify new growth business areas for Dow to study. The Project 2025 reports will be reviewed later this year to find additional business opportunities to study.

On a different line, the results of a proprietary study for Dow to identify potential areas for further study reported on 15 possible business areas. This led to an afternoon of focused brainstorming in these areas. The results of the brainstorming and the summary reports were distributed to the participants and the Ventures Leadership Team. Some of the ideas are being integrated into our formal process for opportunity assessment or are being used as support data for projects which are already underway.

About two years ago, we conducted two other future-based issue analyses. One was a survey of literature from which we extracted pertinent trends or possible events which could affect current Dow businesses or may create an opportunity for a new business. This work was written in a report and distributed to top Ventures and current business management.

The second study was a survey of several top managers in all functions and all geographical areas within Dow. We asked them to work with their staffs to list the most important technical, political, and social trends or issues which would affect their current business or create new opportunities. This report was then recirculated back to them after the data was compiled. We used this as an input into our search for new business opportunities, and presumably they used the results in their strategy development.

In all of these studies we have sponsored or conducted in the past three years, we have found considerable consistency. I believe that the information we collect in these processes is being better used today than ever before and is having a profound effect on our new business development programs. We expect to continue to conduct future-based activities to keep business management aware of trends and events which may affect their areas.

Action Implications

To sum up with some of the operational implications of our look at the last quarter century of scientific and technological forecasting, we suggest actions that would be appropriate for government, trade associations, large corporations, users or consumers of forecasts, and for the think-tank and academic community.

- Almost every field would profit from upgrading its skills and commitments by sponsoring its own forecasts and by orienting its members, whether professional or business, to the value of forecasting.
- Forecasts, to a striking degree, have an amateurish element to them. Key components of an effective forecast are often ignored. Among these components are scientific and technological assumptions, economic, social, and political assumptions, the time frame of the forecasts, method or techniques used to generate the forecasts. We found surprisingly little application of such standard tools as cross-impact analysis or scenarios. The distinction between extrapolative and normative forecasts is often blurred.
- The formal quantitative tools of forecasting are terribly under-used.
- There is almost no critical review of forecasts anywhere. It may be a combination of

politeness or indifference, but the absence of critical feedback on forecasts surely cannot be good for either the field or for the practice.

- There is strong value in bringing outsiders into a forecasting activity in order to avoid the risks of group-think of the insiders talking to themselves.
- Discontinuities, that is sharp disruptions in trends, unexpected events, whether for the good or bad, are a prominently neglected area in the forecasts that we reviewed.
- In looking at the institutional bases of the people who produced most of the forecasts that we reviewed, we found that few of them reflected the names prominent in the futures field. There seems to be something of an intellectual rift between many professional futurists and the forecasting community. Obviously, bringing those two together would be an enormous benefit to each. Futurists could bring to the game a broader sense of possible developments and a clearer sense of the social, economic, political, and institutional implications. On the other hand, a closer linkage to formal forecasting would surely benefit much of contemporary futurism, which is all too qualitative.
- American forecasts, in contrast to those in Europe, tend to pay too little attention to the social consequences of technological developments. However, throughout European forecasts, there is, if anything, an emphasis in the reverse direction, overbalancing concern and attention to social implications and a relative under treatment of the formal side of technological forecasting.
- As far as the Third World is concerned, encouraging formal forecasting there would have some value in opening up potential research opportunities, but far more important would be better insights into future markets and potential businesses for local development.
- We have a clear need and a tangible market for public service forecasts, that is, forecasts which could relate explicitly to policy-making at local, state, and federal government, and for corporations and the rest of institutional America.
- Formal forecasting has the potential to become an active, lively, and potentially entertaining component of public discussion. We have not found a formal forecast developed and presented for radio or television.

The formation of new technologies

What consequences does a new technology have for society, and what consequences does society have for technology?
In this section an overview of approaches is presented with which the relation between technology and society can be conceptualized. Using these approaches it will be shown why technologies often are so resistant to social wishes to change them, and which can be interfacing points for the social influencing of technological developments.

Introduction

We can ask ourselves the question of how the process of technological development works exactly. Knowledge of this development process and the different relevant factors in it may give us a better insight in the possibilities of influencing the development of technology in the future. This can help an engineer better understand his own position. It can also be useful when we want to steer a certain course in developing technology, enabling us to reach specific goals or to avoid negative effects.
Within several scientific disciplines interests toward innovations and technological development exist. After the Second World War these interests had a strongly economic orientation. This stemmed from the problems economists encountered when attempting to explain the post-war economic growth: it wasn't explicable using macroeconomics alone. The economic influence of technology was merely taken into account in the form of a constant. Technological development, however, turned out to be an essential factor in economic growth.
The relations between fundamental research, development and technical implementation took a central role in new models devised by economists. Inventors were only interesting to them if they marketed their inventions themselves. The focus was more on the businessman rather than the inventor. The division between *invention* and *innovation* also stems from this. An invention only became an innovation after it was marketed. This fit into a tradition inspired by J.A. Schumpeter, an economist from the first half of the twentieth century. Inventor-cum-businessmen like Thomas Edison and Alexander Bell became very famous for that reason.

This type of innovation studies led to a fairly linear view on technological development, in which the focus was drawn to the great successes. Figure 9.1 is an example of such a view.

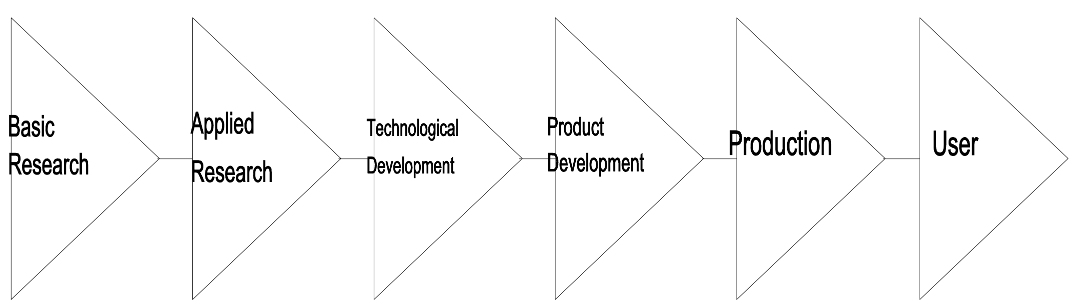


Figure 9.1 A six-phase model of an innovation process

Similar models came under heavy criticism starting in the seventies. The development of technology isn't a linear process, determined only by scientists and technologists, at all. It is more of a process in which choices are continuously made between alternative possibilities of development. Social factors play an important role herein.

This is why researchers tried to form a more expansive framework encompassing innovations, inventors and businesspeople, where the influence society has on the realization of technological developments is given a lot of attention. An important impulse here comes from the history of technology. This has a long tradition of research into the development of instruments or technical systems. Globally, the history of technology has centered itself around both an internal and an external approach. The internal approach often concerns itself with the precise tracking of the development of a specific machine, for example the steam engine.

The external approach places the accent on the influence of the social environment on technology. The development of technology is a resulting variable in this approach, which depends on factors in the social environment.

> **Internal** approach towards history of technology: technology possesses dynamics of its own. Influences from the social environment merely cause ripples in the continuous flow of technological development.
> **External** approach: technological change is completely driven by external forces. Every 'autonomous' development is an illusion as the stability is caused by stabile social conditions.

To get a hold on the process of technological development, various theories and models have been developed. Older forms of technology studies are characterized by a 'black box' approach: Economists and sociologists showed the tendency to view technological development only from the outside, and regarded the technology itself as a black box. Choices made in the technological development process stayed out of the picture this way. Historians did try to expose these choices: they entered the black box but themselves spent little attention on the influence of the environment on technological development. Theories that were developed aimed at explaining technology change as an autonomous process. More recently, visions have been developed which attempt to incorporate social factors into the process of technological development. One of these is the so-called SCOT model. This model emphasizes the possibility of choice in processes of technology change and the interest social factions have in it.

In paragraph 9.2, some theorizing regarding autonomous technology will be briefly explained. In 9.3 we will explain the SCOT model. The development of the bicycle is taken as an example there, because the model was initially designed for that situation. In paragraph 9.4 another model for technological development will be treated; the so-called system model.

Autonomous Technology Development and Technological Determinism

In this chapter we will deal with the question what drives technological changes. An implicit view that is often implicit in popular media is the view that technological change is autonomous. This means that technological change is not influenced through social (economic, social and legal) powers. 'The progress of the technology cannot be halted', or 'As Einstein had not invented the general theory of relativity, someone else would have done it'. Core of this way of reasoning is the assumption that technology is fed by scientific knowledge. Science is only growing, and therefore technology will only become 'better'.

Moreover, improved technologies help to improve other technologies. Hence, technology development is autonomous. This autonomous technology is the main driving force of progress. The state of society is determined by technology. One of the best-known philosophers that approached technology from a deterministic (or even fatalistic) viewpoint was Jacques Ellul. Ellul's constant theme in all his publications is the imminent 'technological tyranny over mankind'. Ellul creates a sharp divide between the classic (Medieval) technology and modern technology. Traditional technology was according to him:

- Limited in its application (because technology had been made for specific functions on a specific place)
- Only marginally dependent on resources and especially dependent on craftsmanship
- Local in its character, (because local circumstances are used, and local culture has to be taken into account).
- The result was that classic technology allowed the possibility of choice, that is to say individuals and local communities could to a far extent determine the shape of the technology that they applied.

Contrasting to traditional technology, Ellul characterized modern technology through:

- Automatism, i.e. there is only one 'best' way to solve a particular problem, which is compelling where ever you are on this planet.
- Self-replication, i.e. new technology strengthens the growth of other technologies. The result is exponential growth.
- Indivisibility. In order to participate in modern society, the technological lifestyle must be accepted completely, with its good and bad sides.
- Cohesion, i.e. technologies of different areas have much in common
- Universalism, i.e. technology is geographically as well as qualitatively omnipresent.

For Ellul this meant that modern technology is devastating for human freedom. In his view, the future of mankind is extremely gloomy, for there is no way back.

Elluls' arguments partly can be recognized in the so-called Unabomber Manifesto. Unabomber, the pseudonym of the Californian mathematics professor Kaczynski, committed in the eighties and early nineties attacks on research institutions and airlines. A characteristic line of reasoning in his manifesto is paragraph 127.

> 127. A technological advance that appears not to threaten freedom often turns out to threaten it very seriously later on. For example, consider motorized transport. A walking man formerly could go where he pleased, go at his own pace without observing any traffic regulations, and was independent of technological support-systems. When motor vehicles were introduced they appeared to increase man's freedom. They took no freedom away from the walking man, no one had to have an automobile if he didn't want one, and anyone who did choose to buy an automobile could travel much faster than the walking man. But the introduction of motorized transport soon changed society in such a way as to restrict greatly man's freedom of locomotion. When automobiles became numerous, it became necessary to regulate their use extensively. In a car, especially in densely populated areas, one cannot just go where one likes at one's own pace one's movement is governed by the flow of traffic and by various traffic laws. One is tied down by various obligations: license requirements, driver test, renewing registration, insurance, maintenance required for safety, monthly payments on purchase price. Moreover, the use of motorized transport is no longer optional. Since the introduction of motorized transport the arrangement of our cities has changed in such a way that the majority of people no longer live within walking distance of their place of employment, shopping areas and recreational opportunities, so that they HAVE TO depend on the automobile for transportation. Or else they must use public transportation, in which case they have even less

> control over their own movement than when driving a car. Even the walker's freedom is now greatly restricted. In the city he continually has to stop and wait for traffic lights that are designed mainly to serve auto traffic. In the country, motor traffic makes it dangerous and unpleasant to walk along the highway. (Note the important point we have illustrated with the case of motorized transport: When a new item of technology is introduced as an option that an individual can accept or not as he chooses, it does not necessarily REMAIN optional. In many cases the new technology changes society in such a way that people eventually find themselves FORCED to use it.)

Besides these fatalistic views, there are also very optimistic technological determinist views. Especially a number of futurologists propagate bright images of future technologies. Unimaginable speeds of transport, the conquest of space as the ‘final frontier’, living at the ocean floor or on Mars, it can all be done. Whether society really needs these techniques is of no concern. It is imaged as the inevitable ‘progress’.
The technological deterministic worldview is dubious:

- It supposes one-way traffic between science, technology and society. Technology is the product is of scientific growth and technological self-replication. However, historical this is incorrect: technology often precedes the formulation of underlying scientific principles. This holds for instance for the steam engine that was already a century in use before Sadi Carnot formulated the Carnot cycle in 1824. The Carnot cycle explains the transformation of heat in work. The first airplane flew in 1903, but the principles of flying by wings could only be understood by the work of Prandtl around 1920.
- Historical analyses show that technological innovation is not a process that will inevitably lead to an optimal result. Choices can be made by social groups.

For technological determinists every technological change has its inevitable course. Actors, (even scientists and technologists that want to change their own work e.g. based on moral or political convictions) produce in the eyes of determinists only a light ‘ripple’ that can be neglected. Although technological determinism is not very fashionable these days, it cannot be denied that there is a core of truth in it: in our globalizing society, there is very little scope for national authorities to influence or even steer processes of technological change.

The Social Construction of Technology

A more recent approach is the SCOT model: the "Social Construction Of Technology". In the SCOT model, technologies are considered social constructions, to which various groups of people have given shape. Central to this view is the notion that people influence the development of a technology by the meaning they attribute to it (this will be explained further on).
The SCOT model has been developed according to a number of case studies – among them the transistor, Bakelite, the neon light and the bicycle. We will explain the model according to this last example.

> **Social Construction Of Technology**: technologies are social constructions, to which different groups of people have contributed. Groups of people attribute different meanings to technologies, which lead them to different perceptions of problems. They influence the influence the construction of technologies by seeking solutions for their problems.

When we consider the development of the bicycle in the traditional linear way, we see the current safety bike (Lawson's Bicyclette) as the final product of an evolution which started as the Walking Machine and where the so-called Penny Farthing represents a transitional stadium.

The gripe with such a representation is, that it isn't clear that choices are made. In this way it appears that the Penny-farthing stays in the picture for too long, even when the safety bike was already introduced – while it was technically inferior to the latter. Moreover, such an

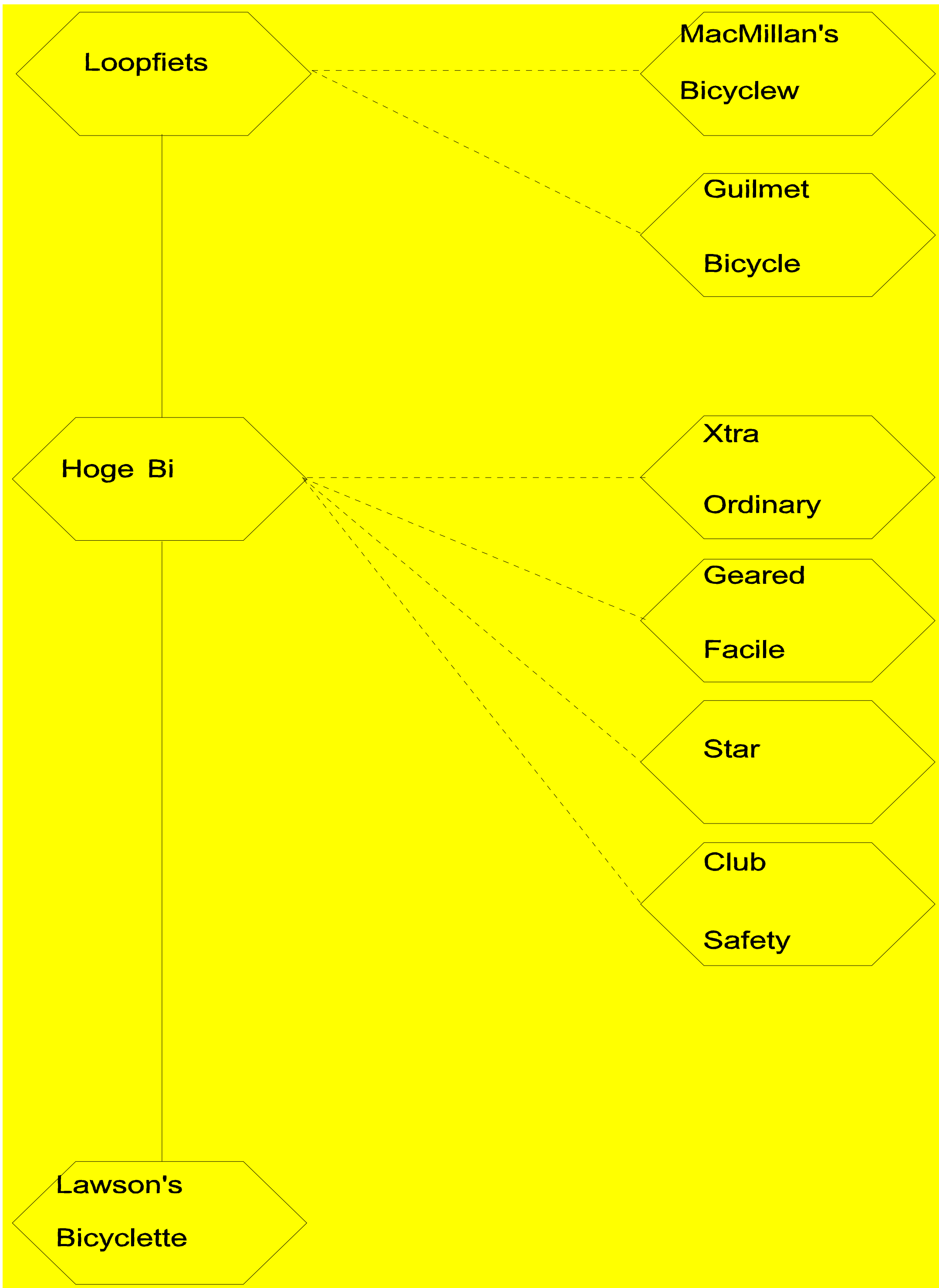


Figure 2. A linear development model of the bicycle.

approach often leads towards the emphasizing of the models, which came out as the winners. One of the goals of the SCOT model is to indicate that choices are being made. The line from the Penny-Farthing to the Safety Bike is a theoretical construction, which doesn't do reality justice, in which the two types co-existed and both were under development. Let us take a closer look at this.

The Penny-Farthing had a large forward wheel and a little rear wheel. Pedals attached to the forward axis propelled the bicycle. Because of the reachability of the handlebars and the turning capabilities of this bicycle the cyclist had to sit almost straight above the forward wheel. The bicycle was fast and efficient, but highly unstable. It was introduced in 1870 and lasted until the end of the century. Lawson's Byciclette was characterized by rear-wheel transmission. This safety bike stems from 1879. By the end of the century it had crystallized into the bicycle, as we know it today. The typical characteristics were, besides rear-wheel transmission, wheels of the same size and air-inflated tires – which majorly contributed to its safety.

To better understand the development of the bicycle, we begin with an introduction to two important concepts in the SCOT model:

- Artifact: a consciously man-made, artificial object.
- Relevant social group: people who are involved in a certain technical development and who hold the same view regarding an artifact.

Around each artifact a number of relevant social groups can be distinguished. People involved with an artifact all hold a certain image of it: they assign a certain meaning to it. Especially important is what people find problematic about the artifact. Groups form (or can be constructed) based on a shared assignment of meaning and, embedded therein, a shared perception of a problem. A group can have different problems regarding to an artifact. When schematized the situation looks like this:

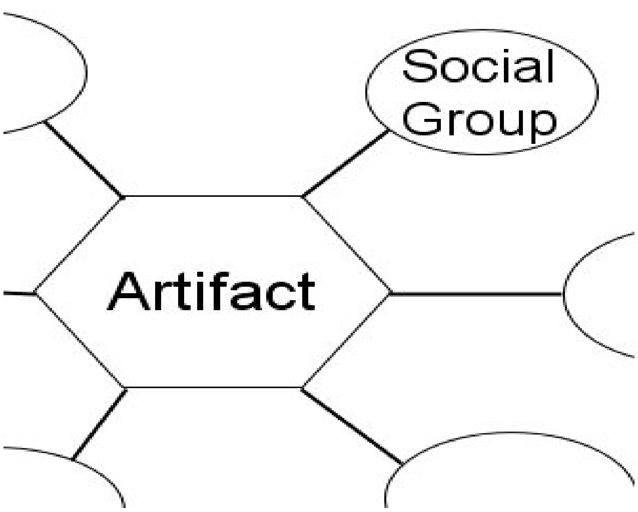


Figure 3. Relevant social groups connected to an artifact.

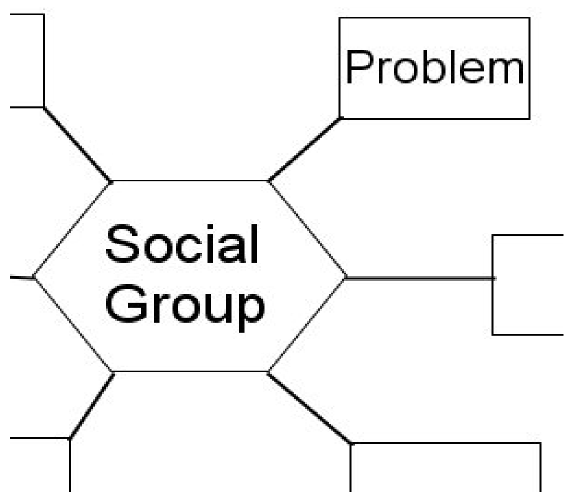


Figure 4. Problems a social group perceives with an artifact.

For problems (or clusters of problems) multiple solutions are conceivable. Through their way of defining the problem, possibly including their solution(s), the different social groups exert influence on the development.
Using the cluster "Technical artifact – relevant social group – problem – solution" we are able to explain the development process of an artifact. An important term here is flexibility of meaning: different social groups attribute different meanings to one technical artifact. This notion is essential to understand for instance the development of the bicycle.

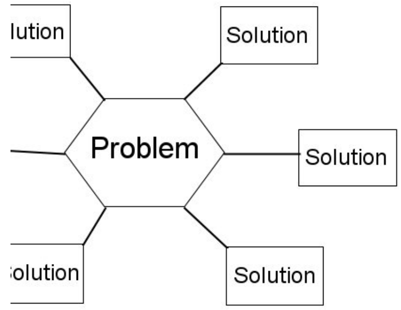


Figure 5. Solutions to a problem

Relevant social groups surrounding the artifact 'bicycle' are the producers and the users of the bicycle. However, we should also involve the 'anti-cyclists' in the story. The bicycle also evoked resistance with people: *"....but when to words are added deeds, and stones are thrown, sticks thrust into wheels, or caps hurled into the machinery, the picture has a different aspect"* (cited in: Bijker, 1984). These anti-cyclists not only had decency problems with female cyclists, but also with the dangers that came with cycling. In London for instance cyclists used wooden sidewalks, because the roads were otherwise unpaved. This evoked resistance with the local population, further enhanced by existing class differences. The largest user group with the Penny-Farthing turned out to be young men of reasonable wealth, which possessed the courage and dexterity to handle the bikes. Besides them was a group of potential users. The Penny-Farthing-riders – young brave and from the higher circles of society, radiated superiority towards their walking or horse riding brethren. For them, the Penny-Farthing was a 'macho machine'. For potential users such as women, long-distance cyclists or older gentlemen the Penny-Farthing was rather considered an unsafe machine.

artifact: a consciously manufactured man-made object.
relevant social group: a group, which assigns a specific meaning to an artifact.
flexibility of meaning: different social groups assign different meanings to the same artifact.

Because the cyclist was seated almost directly over the middle of the forward wheel, with his/her legs far from the ground, every stop or bump provided the risk of falling over. These different attributions of meaning also spawned different directions of development:

1. For the sportsmanlike Penny-Farthing rider enlargement of the forward wheel was the best way to increase speed. This culminated in 1892 in the Rudge Ordinary with a forward wheel diameter of about 1.4 meters. The fact that this only made the bicycle more dangerous was considered by the specific user base to be more of an advantage than a drawback.
2. To make the unsafe Penny-Farthing suitable for other users, experiments were done with various different models. The wheels were reversed, such as with the pony star, or the forward wheel was made smaller and the saddle was put further backwards as in Lawson's Bicyclette (When the safety bike finally developed into a faster bike than the Penny Farthing, the latter's fate was sealed).

When we regard the technical development of the bicycle using the SCOT model, it yields figure 9.6. This image shows that the application of the SCOT model results in a different view on the development of technology, in this case the bicycle, from a traditional phase model.

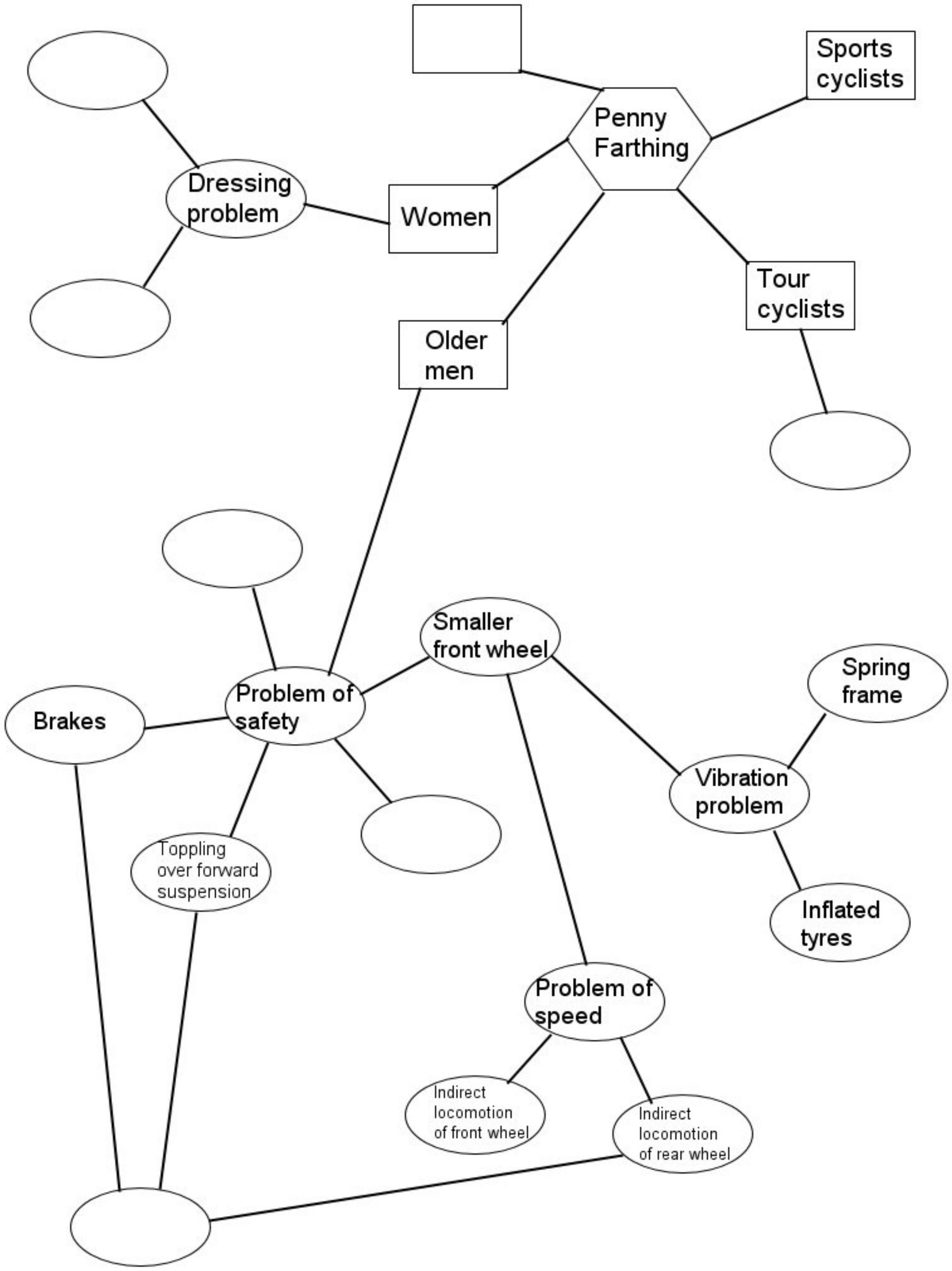


Figure 9.6. The SCOT model applied to the development of the bicycle.

In the SCOT model the assignment of meaning, the signaling of problems and solving them are aspects of the process of technological development. It is entirely possible that the final product as a result hasn't been foreseen by any of the parties involved. This certainly holds for the safety bike. The formation of it can't be clearly dated, but stretches over a number of decades. At the start of the process that led to this bicycle, none of the involved had a clear vision of it in its final form. Several different models were present at that time too, among others Lawson's Bicyclette, the Kangaroo and the Facile.
Paraphrasing, the development of an artifact globally consists of three stages:

1- interaction between technological development and social groups
2- variance in problems and solutions
3- the choice for one solution

Of course, the solution doesn't need to be of a permanent character: the solutions obtained are temporary forms of stabilization. What is used as a solution doesn't have to be an example of brilliant engineering. In the case of the Penny-Farthing for instance, users complained that the handlebars would be in the way if the rider fell over forward as a result of a sudden stop. The solution that was proposed for this problem was to make the handlebars detachable in such a case to enable the rider to land on his feet. Although such examples of solutions couldn't be dismissed outright, the solution didn't last long in this case.
When explaining the development of an artifact the term 'technical framework' is also important. With 'technical framework' is meant all of the solution strategies, theories, skills, user practices, goals, values, ethical standards, and ethics regarding a certain technology of a certain social group. This determines the thinking, acting and interaction within the group. To show the involvement of a stakeholder within a technical framework the term 'inclusion' is used. Inclusion in different technical frameworks at the same time is possible. An electrical engineer has a high inclusion in the technical framework connected to his own discipline, but he may also have followed subjects at the faculty of physics during his course. This makes it possible for the electrical engineer to be involved in another technical framework too.
The technical framework can be used to explain why a certain solution is chosen. In the case of the Penny-Farthing, it's easy to picture that the builders and users only had one goal in mind: increasing the bike's maximum speed. Caught in the technical pattern this bike adhered that was only possible by increasing the size of the forward wheel.

Technical framework: a set of theories, skills, user practices, goals, values and ethical standards regarding a specific technology.
Inclusion: the degree to which a stakeholder is involved in a certain technical framework.

In the above the concepts central to the SCOT model have been introduced and illustrated according to the example of the bicycle. Using the model we can explain why the Penny-Farthing was such a success for a long time, even when the safety bike had been introduced. The key to this pattern lies in the 'macho machine' function the Penny-Farthing had.
The previous analysis also offers a starting point in case we want to steer the development of a technology. It means that first it should be studied what a technology will mean to the different people involved, and the relative importance of their opinions. The process of choices made won't be completely predictable – the steering of technology will therefore have to be a continuous process. Steering (government-) agencies have to be aware of the (sometimes hidden) meanings a technology can obtain, and adapt their steering activities accordingly.

Technological development as system development[17]

In the social-constructivist (9.3) and the evolutionary-economic (10.4) views of technological development the development of individual artifacts takes a central role. Thomas Hughes posed another, opposing vision to this in 1983: not so much the development of individual artifacts matters, as much as the formation and further development of technological systems. Although he used an obvious example to illustrate his view, electrical systems, he held a broader view: The basis of the system-wide approach to technological development is the assumption that development of all large-scale technology (not only electrical systems) can be studied as a history of developing systems. The history of a specific system, the electrical system, thus only serves as an illustration of a more general view. After those done by Hughes, many other but similar studies were published.

The electrical system

The construction of electricity grids has been an impressive and major event. Not only because of the technical feat it consisted of, or the development of scientific knowledge that was needed to do it, but particularly in the social, economical and political effects of the distribution of electrical energy.

An enormous network of electrical wiring organizes the way we live. Inventors, engineers, managers and entrepreneurs have organized our world by developing this energy network. In the years between 1880 and 1930 the most important decisions were made, and the technology was developed for this network.

Therefore, by studying this period the ordering, integrating, coordinating and organizing of this network and the society, which it is part of can be analyzed. Electrical energy systems require a sense of efficient action, the ability to make rational analyses and the ability to effectively deal with 'vague' economical, political and social developments from their inventors, operators and managers. Leading engineers have acknowledged that their desire to 'clear up' matters often has to be moderated to accommodate for the 'disorganized' phenomena that make our society so vital.

How did the small-scale city lighting systems in the eighties of the 19th century evolve to the regional electricity companies of the twenties? The problem here is explaining the change in configuration of electricity-producing systems during the 1880-1930 period. These changes can be pictured in a series of network diagrams. However, energy systems are cultural artifacts. This is the reason why an explanation of these changes has to incorporate several different areas of human activity such as technology, science, economics, politics and organizational science.

Electrical energy systems embody the physical, intellectual and symbolical resources of the society, which produces them. Therefore change in energy systems can't be viewed separately from changes in available resources and the aspirations of organizations, groups and individuals. Systems for the production of electrical energy, which arise in other societies and in other eras, often hold a number of basic elements in common. Variations in these basic elements however often occur, too. These stem from variations in the availability of resources, difference in traditions, different political situations and different economical practices. This is why electrical energy systems are both a cause and an effect of social

[17] This excerpt is largely based on the introduction to NETWORKS OF POWER, Electrification in Western Society 1880-1930, Thomas P**. Hughes, 1983**, pp. 1-17.

changes. Energy systems reflect their environment, and change it too. They possess internal dynamics of their own too, however.

Technological systems

> A **technological system** consists of linked components, which are centrally controlled to a certain extent. The system boundaries stem from the reach this central control has. The goal of this central control is to optimize the behavior of the system regarding a certain goal.

How can technological systems, which make up an increasing part of our environment, be defined? Ludwig von Bertalanffy, a well-known systems theorist, needed a complete book to define 'system'[18]. Therefore a more inadequate description of the system concept will have to suffice here. However, several characteristics a system has are very common. A system consists of components or parts, which are related to each other. These components are connected by a network, or structure, which may even be more interesting than the components themselves. The linked components of a system are often centrally controlled. The system boundaries follow from the range of this central control. Control is present to optimize the performance of the system as a whole, and to direct the system towards a common goal. The goal of an electricity production system, for instance, is to convert the available energy, input, to the required output, which is fulfilling the demand for electricity. Because the components are linked to each other, the state or activity of one components influence those of another.

The network of connections between these components determines the system configuration. A system can for example consist of horizontally, star-shaped or vertically configured components. A horizontally arranged system connects components, which perform similar functions, although not necessarily of the same scale. A vertically connected system connects components in a functional chain. In this way, an electrical system of the horizontal kind connects power plants, regulated by a central office. A production system of the vertical kind for instance connects a coalmine to a power plant under central management to take care of the attuning of coal production to the demand for power. Systems are designed hierarchically too, whereby smaller, in a sense independent, systems contribute to the control of the larger, encompassing system.

Systems also interact with each other. Those parts of the world, which don't belong to the system but do influence it are collectively called the 'environment' of the system. Parts of the environment can sometimes be made part of the system by bringing them under the system's control. An open system is subject to environmental influences; a closed system is not. Hence, the behavior of a closed system is, in principle, completely predictable – that of an open system isn't.

Some systems are planned completely (from the start), such as the construction of a polder, while other systems grow more incrementally, possibly merge or split, etc. The term 'system' here refers to a technical system, like an electrical transmission system. However, not necessarily all components of the system must be technical in nature. Maintenance services, training institutes, administration etc. are as important to the system as its technical components.

Electricity-providing systems consist of electricity production, transformation, control,

[18] Ludwig von Bertalanffy, 1968, General System Theory, Foundations, Development, Applications, New York.

transmission networks, distribution networks, and user components. In the period of 1880-1930 the production of electricity was done using steam engines and steam- and water turbines. Different generator types were coupled to these ‘prime movers’. Transformers slowly grew into the main way of controlling the characteristics of the electrical current in transmission and distribution. User components were lamps, engines (both stationary and traction used in trams and trains), heating and electrochemical equipment. The system served widely different purposes. Transmission distances in this period increased from a couple of hundred meters to regions covering tens of thousands of square kilometers. Distribution networks transported the electricity from the transmission network to industries and homes. Control components regulated the electricity production system to keep characteristics such as voltage and frequency at the right level, and made sure the system functioned optimally regarding its goals such as efficiency, profit generation, reliability, etc. The components most difficult to define are those at both ends – both on the side of demand and the side of supply. For instance, are the mechanical ‘prime movers’ a part of the electricity system? Waterpower sometimes escapes control by the system. Are the different usage intensities a part of the system, when you consider the fact that the grid sometimes has influence on them (peak loads) and sometimes not?

Hughes chose to take the ‘prime movers’ into the definition of the system because inventors, engineers and researchers treated them as such and because they were mainly controlled by the system. However, the definition of his system boundaries remains somewhat unclear. The invention of electric engines was for an important part guided by the characteristics of the electricity system at that time. The functioning of those engines however can be controlled only in a very limited way so they don’t form part of that system.

Phases in system development

Hughes analyzes electricity systems, which were formed in different places (New York, Chicago, London, Berlin, California) and at different times. Yet they nevertheless are connected in his view, because of the fact that they all behave according to the same model for evolving systems. In that model for system development different phases can be distinguished: and in different phases of system development different characteristics hold a dominant position. Moreover the model indicates the skills managers must possess in each of the phases, and the guiding interests.

In the first phase the emphasis is on the invention and development of a system. The professionals dominating this phase are inventors/entrepreneurs, who differ from regular inventors by their attempt to organize the entire process from invention to ‘ready for use’. Edison, of course, is the supreme example of such a person. Engineers, managers and banks are also important during this phase, but they are of less importance than the inventor/entrepreneur.

In the second phase the most important process is technology transfer from region to region or from continent to continent. The transfer of Edison’s electrical system from New York to Berlin and London is an example of this. During this phase various groups are involved in the development, such as the inventor/entrepreneur, traditional entrepreneurs and banks.

The essential characteristic of the **third phase** of the model is system growth. Growth of systems is analyzed by means of the concepts ‘reverse salient’ and ‘critical problem’.

A **reverse salient** is that part (or parts) of a technological system, which is lagging behind in development and therefore limit the growth of the system as a whole.
A **critical problem** is a redefinition of the reverse salient into a (principally) solvable technical challenge.

Because system components often grow at different rates, parts of the system can be identified which lag behind the others in growth and are limiting the growth of the system as a whole. The term ‘reverse salient’ stems from military history. In that context the term signifies a section of a front line, which is lagging behind in the advance. This metaphor is well chosen, because a progressing military battlefront also often shows irregular and unpredictable behavior, just as a developing technological system does. Often, military people will direct all their efforts towards fixing a ‘reverse salient’. The same goes for the development of a technological system.
Inventors, engineers, entrepreneurs and others direct their creative and constructive forces mainly at the correction of ‘reverse salients’, in such a way that the system functions optimally in fulfilling its tasks.
If the reverse salients are identified, they will often be translated into a series of critical problems. The redefinition of reverse salients as a series of critical problems is the essence of the creative technological process in the system: An inventor, engineer or scientist transforms an amorphous challenge (the lagging behind of (parts of) the system) into a series of problems, which are expected to be solvable. This is an essential part of the engineering profession: being able to redefine unstructured problems into a series of solvable, critical problems. The confidence in the solvability of the reverse salient increases dramatically after it has been turned into a series of well-posed critical problems. Correct articulation of problems usually helps a lot in approaching a solution. If engineers are capable of correcting the reverse salient in this way, it usually leads to growth of the system. However, sometimes a situation occurs where a critical problem appears to be unsolvable.

Example

Around 1880 Edison realized the first electrical system in the world on Manhattan. This system was based on the distribution of direct current. The distribution of electricity only took place over short distances (hundreds of meters), using large and massive copper wiring, at low voltage.
This system experienced as its most important reverse salient in its later stages of growth the fact that it was only economical to use in built-up and localized areas, because of the high transmission losses. This reverse salient could in principle be translated into different critical problems:

- reduction of transmission losses
- improving the efficiency of smaller power plants
- battery-based distribution in sparsely populated areas
- ?

The reduction of transmission losses was the most important problem to which this reverse salient led in practice. Despite an accurate definition of the problem the inventors and engineers using direct current at the end of the 19th century weren’t able to find a solution to this problem. In the end other inventors found a solution outside of the DC system, to be specific: alternating current, which enabled the system to transform the voltage easily by transformers. This caused the coexistence of two conflicting systems for some time. By applying all sorts of transformation technologies these systems could coexist for some time, until the AC system became the dominant one. New systems can hence form where old systems are unable to resolve their reverse salients by their own means.

When a system grows, it gains momentum. The **fourth phase** of a system is characterized by substantial momentum. A system with a substantial momentum possesses mass, velocity and a direction of motion. The mass consists of machines, equipment, structures and other physical artifacts into which capital has been invested. The mass also stems from the involvement of people who possess professional skills specifically suited to the system. Entrepreneurial businesses, government services, trade unions, educational institutions and other organizations, which are directly attuned to the system's core also contribute to the momentum of the system. Taken together these organizations form the culture of the system. A system also has a measurable speed of growth. That speed often increases in this phase. A system also possesses a direction of motion, i.e. goals. The definition of clear goals is more important for a new system than it is for an older one. In older systems the momentum gained gives the system inertia in its further development.

In electrical systems the public and private utilities were the institutions controlling the system. From 1890 onwards to the First World War the most important utilities in the USA, Germany and England concentrated on the supply of electricity to densely populated and industrialized urban areas. The decisions made by their managers in this period were of more importance to the character of the systems than the decisions made by their engineers and inventors. In this phase the conflicts between the utilities and other social stakeholders were often of great interest.

Despite the momentum of systems and their inertia there are more or less coincidental factors pushing them in other directions. The First World War for instance made the supplying companies direct their activities more towards fulfilling the energy demand generated by the industry rather than a maximization of profit. This often required a cooperative stance and formed structures, which would persist later on. External forces can hence bring about a reorientation of system goals, even when systems have already gained a lot of momentum.

The **last phase** of system development can be characterized by the qualitative change in the nature of the occurring reverse salients, and by the advent of financers and consultants as problem solvers. Managers played the leading part in the momentum-gaining phase. In the later phase, which concerned itself with planned and developing regional systems the most important reverse salients were problems stemming from the need for the financing of large-scale systems and the clearing up of legal and political barriers. Financers and advising engineers were able to respond adequately to these problems.

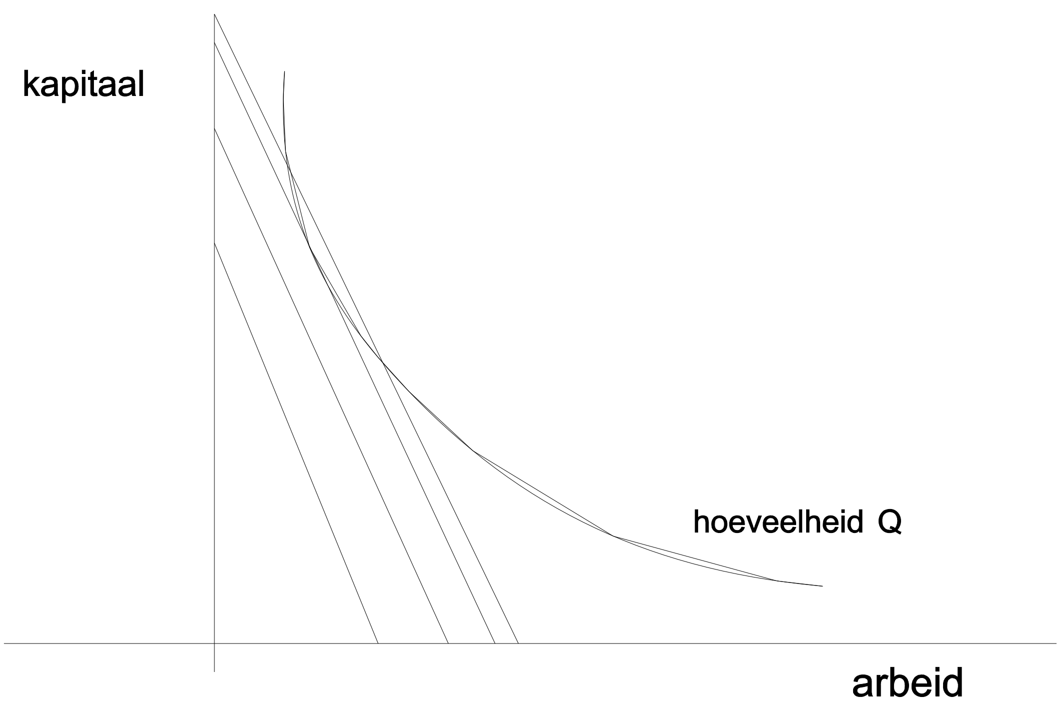
This phase was also characterized by a continually increasing competence, mostly present with advising engineers and managers, in the effective planning of systems.

Economic approaches to the formation of new technology

What consequences does a new technology have on society and what consequences do social developments have on technology?
In this section an overview will be presented of the way economists think about technology. Attention will mostly be paid to the (quasi-) evolutionary approach. Finally the phenomenon of 'path dependence', where society does end up with the most optimal technologies, will be analyzed.

The neoclassical economic framework

Neoclassical economic theory has been around for quite some time now. It formed in the closing years of the 19th century and still holds great influence today. Neoclassical theory plays a major part in government planning for instance in the models used by the CPB, the Netherlands Bureau for Economic Policy Analysis. We will only treat the elements of this theory, which serve our purpose here.
Neoclassical theory assumes entrepreneurs wishing to produce a certain quantity of a product have a choice between different combinations of production factors. The term 'production factors', when used in economics, signifies all that is needed for production: manpower, monetary resources and physical resources. It can be decided to perform production in a capital-intensive way (using lots of machines) or in a labor-intensive way. Neoclassical theory assumes this choice primarily depends on the cost of capital (interest) and manpower (wages).



The production function

> **The production function** indicates with what distributions of capital and manpower a certain volume of a product can be produced.

This choice is pictured in the production function (see figure). The function yields all combinations of manpower (X-axis) and capital (Y-axis) with which a certain quantity Q of a product can be produced.
The entrepreneur is hence free to choose any point on this curve. Where on the curve he decides to be, follows from the pricing of manpower and capital. When certain prices for manpower and capital are given, each of the slanted lines in the figure show the amounts of capital and manpower which can be bought for a certain, fixed price. For each line more towards the right (or towards the top) this price is higher. The entrepreneur who wishes to produce his amount of products Q at the lowest possible cost, chooses the point on the curve which lies tangent to the cost function – this is the point on the curve where the total costs of manpower and capital combined are lowest.
When the pricing of manpower and capital changes, the slanted lines change slope and hence displace the tangent point. This encourages the entrepreneur to switch to a different point on the curve, i.e. to a different distribution of manpower and capital. If for example manpower gets more expensive, the lines will be a at a steeper slope and the entrepreneur moves to a point 'higher' up on the curve, to a more capital-intensive production method. Of course this doesn't happen instantly, but at a time, which suits the entrepreneur – for instance, when the machines need to be replaced anyway.
Another interesting element of neoclassical theory is the set of supply- and demand functions of a product. The theory assumes that supply and demand depend on the price of a product. The higher the price, the larger the supply but the lower the demand (figure 10.2). If the product is to be marketable, the supply- and demand curves have to intersect somewhere, lest the product become a failure. This point of intersection determines the price of the product and the quantity sold. Because of external influences the supply- and demand functions can change forms. By applying cost-saving measures the supply curve can be made to take a lower position (shifting to A'). This makes the point of intersection with the demand curve shift from a to b, which means a greater quantity is sold at a lower price. Because of an increase of consumer income, the demand curve can shift upwards (to V'). This also means a greater quantity of the product will be sold, but at a higher price this time (point c).

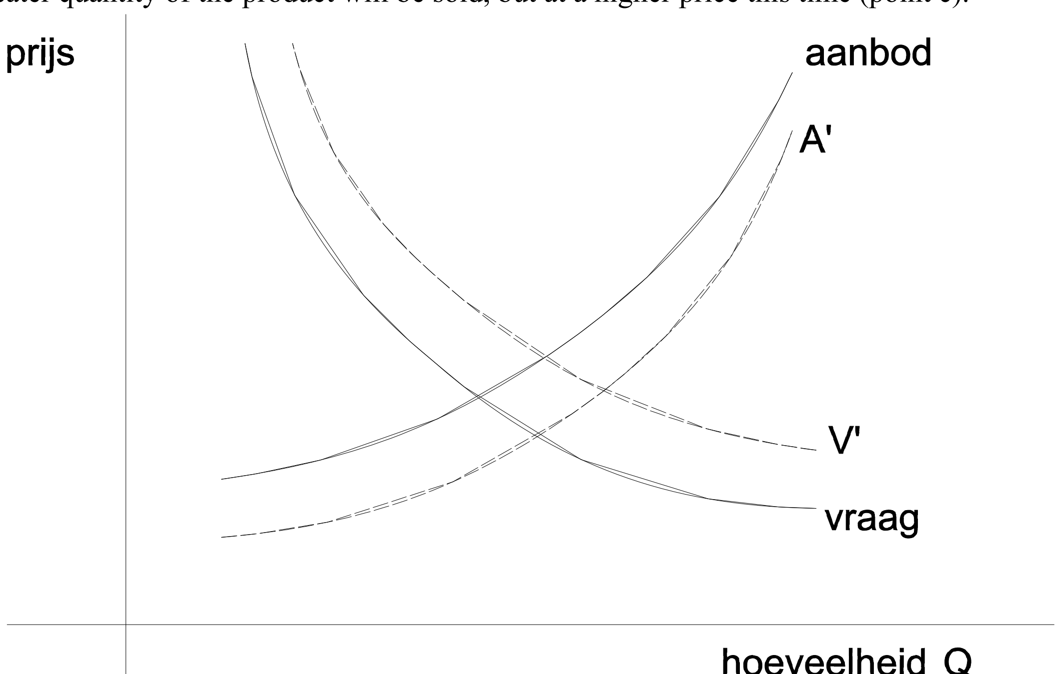


Figure 2 Functions of supply and demand

For our purposes neoclassical theory offers some possibilities, but also some important restrictions. The possibilities lie in the fact that the influence of factor costs on technological changes can be determined. After all, every choice made between capital and manpower entails a technological choice, a choice between production technologies. Using this theory we can hence gain some understanding about the influence of manpower costs and interest rates on the choice of production technologies. Sometimes the government can even decide to directly influence factor costs to influence the outcome of technological choices. For instance, the government can impose taxes on energy use to stimulate businesses to use more energy-efficient production methods. Also taxes on emissions can be used to reach technological changes. These are examples in which the costs of a production factor are artificially enhanced in order to obtain a different technological solution. Even changing consumer behavior patterns can be obtained this way: by enhancing the costs of car use, travelers can be stimulated to make more use of public transport.

Neoclassical theory also has many limitations when it concerns the role technology plays. Its biggest problem is the lack of explanation for technological innovations. Neoclassical theory discerns between the aforementioned shifts along the production function and change in position of (part of) the production function itself. In the first case there is no mention of innovation – existing technology is being used in the new situation. In the second case, innovation takes place. (Part of) the production function shifts towards the origin of the coordinate system. After all, an innovation (in production technology) implies that an entrepreneur can make the same product at a lower cost (in manpower and capital taken together), which also makes the supply function drop and hence benefits the customer in the end.

Where the innovations come from isn't explained at all in this theory. It is assumed they come from outside the economic system (economically exogenous), for instance from technological development itself (technological determinism). However, it is clear this isn't the case: economic factors do certainly have an influence on the direction in which innovations take place. An increase in the pricing of electrical energy for instance won't only stimulate entrepreneurs to search among alternative existing production methods for more efficient ones, but will also provide a stimulus for research into more energy-efficient technologies.

A second problem is the assumption that a large choice of production methods is available, with a large range of costs in production factors. In practice, this is rarely the case. Usually, there are but a few different alternatives for the small part of the production function around which the prices of capital and manpower fluctuate. If those prices start to deviate strongly from that region, all sorts of new technologies will have to be developed. In this way it can be seen why wars, when there appears a shortage of certain production factors, always are a stimulus for the development of new technologies. So, even for movement along the production function technological innovations may be needed.

With this conclusion the difference between movement along the production function curve and movement of the curve itself disappears. In the case of most innovations it isn't clear which of those two changes the innovation belongs to. There are more points of criticism against neoclassical theory, such as:

- Entrepreneurs usually don't research all possible methods of production before they make an 'optimal' choice, but they limit their choice to a few distinct options out of which a 'satisfying' solution is chosen.
- The theory doesn't yield insight into the consequences of changes made to the final product in terms of the production technology used for it.

- The supply- and demand functions are abstractions, which have no directly observable equivalents in the real world. Besides, they are of a short-term nature: if the price increases, entrepreneurs will be stimulated to utilize their production capacity to the fullest and hence the supply will increase. But after a while they will structurally enhance their production capacity and the price will drop again to the previous level. In this way, the supply function doesn't sketch the correct course of events in the long run.

The same goes for the demand function.

Long waves

Kondratieff was a Russian economist who in the 1920-1930 period performed research into long-term waves in the economy. The notion of waves occurring in economy stems from Marx (short term economic cycles theory). Kondratieff specifically focused his attention on long-term waves of economic growth and regression, having a 50-year cycle approximately. Kondratieff saw the irregularity in the replacement of capital goods as the main cause for these cycles. He concluded that the waves were accompanied by irregularities in the appearance of technological innovations. Schumpeter continued this line of thought in the thirties. He was an economist working at Harvard University in the US. Schumpeter[19] perceived (clusters of) innovations as the main cause for economical waves (called 'Kondratieffs' by him). He perceived the following pattern:

> **Long waves**: wave pattern in the development of the economy having a period of approx. 50 years. This is caused by the development of new key technologies.

- 1787-1842 First Kondratieff: cotton, iron, and steam
- 1843-1897 Second Kondratieff: railroads
- 1898-1939 Third Kondratieff: electricity, cars.

Later on others added 'steel' to the third Kondratieff, and identified a fourth Kondratieff: chemistry, electronics, aircraft[20]. In his book from 1939, Schumpeter considered entrepreneurs as being the driving force behind innovations, because they turn inventions into marketable products. The innovations occurred very irregularly in time in his view, and therefore caused sudden breakthroughs in economy. Later, Schumpeter described the occurrence of innovations as being more spread-out[21]. Not entrepreneurs themselves play the major role, but research laboratories (which is why he tended to support planned socialism more). In research laboratories existing knowledge is being improved upon in an 'evolutionary' manner. This thought got incorporated into the current thinking of evolutionary economists.

'Push-pull' debate

---

[19] Joseph A. **Schumpeter**, **1939**, Business Cycles, New York: McGraw-Hill.
[20] Ch. **Freeman**, C. **Perez**, **1988**, in: G. **Dosi** et al (eds.), Technical Change and Economic Theory, Pinter, London, 1988.
[21] Joseph A. Schumpeter, 1942, Capitalism, Socialism and Democracy, RKP.

In this debate among economists the central question is: what is the explanation behind the occurrence of technological innovations? The proponents of the ‘technology push’ theory view the developments within science and technology as the main driving force behind technological innovations, and hence share their side with technological determinism. The ‘demand-pull’ theorists mainly see the market demands as the main cause for technological innovations. Market demand largely determines the formation and introduction of new technological possibilities. Representatives from both camps have supported their theories with a host of empirical research. These schools of thought have closed in on each other however, and a number of economists therefore see the combination of technology push and market pull as the driving force behind technological innovations. A tension between technology push and market pull exists, in which processes of variation and selection led to technological changes. What is new about this view versus technological determinism mainly flows from the role, which economic factors play in the innovation and diffusion of technology.

> **'Technology push'**: developments within science and technology are the main driving force behind technological innovations.
> **'Demand pull'** or **'Market pull'**: demand (or change in demand) from the market causes technological innovation.

The evolutionary theory

While economists expressed their first and important criticism on the notion of the autonomous character of technological development, some others went a step further by pointing towards the shortcomings of a strictly economic explanation of technological development and –introduction. Within the discipline of economics it was mostly Nelson and Winter (1977 and 1982) and Dosi (1982 and 1988) who, building on the ‘push-pull’ theory and Schumpeterian additions to neoclassical economics, attracted attention towards the role which social-cultural and institutional factors play in the processes of innovation and diffusion of technology.

From the science of economics the Schumpeterian models for technological change were further expanded to what are now often known as evolutionary theories. The theory on different technological trajectories by Nelson-Winter/Dosi is perhaps the most articulated evolutionary theory.

The development of technology can, according to Nelson and Winter, be interpreted as an ongoing succession of variation- and selection processes, which are directed towards solving technologically defined problems. New technologies or amendments to existing technologies are constantly invented and selected for usage. These variation- and selection processes don’t just occur ‘at random’ or for no reason, but show a clear structure. A certain rigidity and inertia is present in the rate of change of technology, which prevents

> **Evolutionary approach to technological change**: Development of technology is a succession of variation- and selection processes geared towards the solving of technologically defined problems. A certain rigidity is present in technology, which often only permits small changes in existing technology to occur. However, such a new variation doesn’t always survive the selection by clients, governments or other stakeholders.

variations from cropping up without limits. There is a certain regularity and direction to be found in technological development, which is encapsulated by the concept of 'technological trajectory'. What ensures the control, the structure in the processes of technological development?

Regime and trajectory

Nelson and Winter introduce the phrase 'technological regime' for this, which encompasses the extents of change and development of a technology within a certain problem area. Dosi uses – analogous to the structure found in development of scientific knowledge – the term 'paradigm'. A technological paradigm or technological regime forms the dominant cultural matrix of technology developers and encompasses a limited number of scientific principles, insights and heuristics (searching rules) and a limited number of physical technologies. Central to the technological regime or paradigm are the exemplar and the heuristics. The exemplar is the basic technological design from where all subsequent adaptations and developments are beginning. Starting from this base technology, the direction in which is searched by the processes of variation within the technological trajectory is determined by the heuristics. The development of a technology in a certain problem area takes place over a long time within the boundaries of such a technological regime and is pre-structured as such.

> A **technological paradigm** or **regime** is the way in which technology developers perceive their current and future technologies: it encompasses scientific principles and theories, rules derived from practice, rules for searching solutions for problems and (successful) examples.

> **Technological trajectory**: path of development of a technology governed by a specific regime.

Now can also be described more clearly what a technological trajectory is exactly: the changes in the technology which take place within the framework of such a technological regime or paradigm, in other words the 'direction of progress' within a certain technological regime. A technological trajectory can be seen as the whole of all 'standard' problem-solving activities, which are circumscribed and designed by a technological regime. A technological regime and the direction of development of technologies are to a certain extent inert and subject to little change. Nelson and Winter mention 'natural trajectories', which partly possess dynamics and an optimization pattern of their own. Even when it becomes likely that within the current dominant regime and trajectory the problems which are focused upon can be solved in a less satisfactory way than they could be in a different trajectory or regime, it doesn't mean that the current trajectory is abandoned. Major breakthroughs in the treatment of problems using completely new technologies, in other words changes in the technological regime, occur only sporadically. A change in technological regime almost always encompasses a change in technological trajectory, because different heuristics and base technologies are applied and the dominant cultural matrix, in which the technology developers reside, shifts showing sudden new solution possibilities.
On one hand a technological regime restricts the development of technology by imposing a limitation on the possible variation in new technologies to be developed. On the other hand such a regime enables accelerated development by concentrating efforts and use of resources in one specific direction of research and development. A technological regime dictates how

solutions to problems should be found, and what tools are available to do so.

Selection environment

Which factors influence technology development, in the end? Within evolutionary theory the term ‘selection environment’ is used to indicate the collection of stakeholders, structures and institutions, which determine the nature of the selection process. This selection can take place on three distinct levels: selection of a technological paradigm, selection of a trajectory and selections within a trajectory. This selection environment consists, according to Nelson-Winter and Dosi of three dimensions or types of factors:

1-Science and technology; this pertains to the momentary level of knowledge and technology, the artifacts present, the type and scale of research institutions, the technology present in neighboring fields, etc.;

2-Economy; this regards for instance the prices of factors, market structure defining competition among companies (also internationally), the spread of income, consumer demand, government funding possibilities, structure of production etc.;

3-Social-cultural and political base; this encompasses the balance of power, the distribution of possessions, the legal situation, the cultural matrix of scientists and technologists, government policy and governmental measures, etc.

Which actors, factors and institutions from this selection environment are most important in technological change differs for each development process and can only become apparent using historical case studies.

Using this model a significant number of case studies have been performed, also in Holland. They keep showing that the Nelson-Winter/Dosi model offers enticing anchors and concepts to describe and analyze the development of technology.

Quasi-evolutionary theory

The relationship between technological development and selection environment is a complex one. Influence on the development process originating from the selection environment occurs both ‘ex post’, after technologies have been developed, and ‘ex ante’, before the developments take place. An example if this last case is the anticipation of the increasing demand for ecological products. Moreover, Van den Belt and Rip observed that not only does the selection environment influence the technology development, but the reverse process also occurs. Technological developments pose demands on (and change) the selection environment, for example in those cases where changes in the production process have consequences for the organization of the working environment. Because of these statements the term ‘quasi-evolutionary theory’ is often used nowadays: the selection environment of plays an important role in the selection process, and the selection environment is often strongly influenced by the variation process (both by the result of that process and by the parties involved in it). Van Lente summarized quasi-evolutionary theory in these 7 statements:

Scientific research and technological development activities can be seen as a series of searching processes leading to products like artifacts, texts or skills.

Searching processes are led by searching rules, which hold the promise of being practically successful. They don’t guarantee it, however.

The results of R&D are, in the end, a result of a process of selection and variation.
The combination of searching processes which yield variations, which are selected elsewhere differs from biological evolution because in technological development selection and variation are mutually dependent. This is why the process is called a 'quasi'-evolutionary process.
Expectations are shared and connected to each other to such an extent that they can be defined as a cultural matrix of expectations.
A technological paradigm can be said to be present when a certain coherent set of searching rules has become dominant among producers of variations.
The selection can take the form of a strategic 'game' between players. Actions players take are determined by anticipating other players' actions and on expected future variations.

In conclusion

To what extent does this evolutionary theory definitely separate itself from technological determinism? The development of technology is seen by Nelson and Winter, Dosi and many others who followed in their footsteps as a process determined by more than solely the inescapable logic of science and technology. Although this approach doesn't preach simple technological determinism, it concedes that technology does possess certain dynamics of its own. The development of technology along a technological trajectory within a regime is relatively autonomous, having a local optimum of its own. Within such a trajectory marginal changes under influence of the selection environment can take place. The usage of biological-technical terms such as selection, variation, evolution and natural trajectories reinforces the idea that technological change possesses a certain inherent logic. Moreover the selection environment mostly influences changes within the technological regime, while the influence of the selection environment on the change of regime itself is still unclear.

VHS, MS-DOS, Qwerty, twists of fate[22]

Who still remembers cp/m, Philips-Miller, or Stereo-8? Cp/m was a popular operating system for PCs in the early eighties, Philips-Miller was a recording device from the thirties and Stereo-8 once was a popular cassette recorder in the United States. All three they were pushed out: cp/m by MS-DOS, Philips-Miller by the tape recorder and Stereo-8 by the compact cassette recorder.
Windows, the CD and the 35 mm photo film are products for which it's hard to imagine alternatives. It seems as if there have never been any alternatives to them, as if these products possess a superior quality over any other possibilities. Still, they often are no more than the 'lucky' winners of a competitive battle.
Take MS-DOS, for instance. If the wife of Gary Kildall (cp/m's inventor) hadn't sent away the IBM-managers who rang the doorbell somewhere in July 1980, they probably wouldn't have driven to the Microsoft offices and this article would maybe have been written on a PC running some cp/m version.
The unpredictable process in which one technology remains as the market standard can only partly be explained by economic theory. In many business branches a few large, dominant

[22] This paragraph is taken from Intermediair, May 22nd, 33rd annual, no. 21, pp. 47-51, author: Gerben Bakker

market leaders are present producing the so-called ‘A-brands’. Besides those there usually are tens of other, smaller manufacturers. Products are judged against the ‘standard’ of the market leaders. If they’re more luxurious they’re called ‘premium products’, are they cheaper they’re called ‘discount products’.
With technological products the situation is different. Economists call it the ‘winner takes all’ principle: with technological standards there is one system dominating the market. Sometimes there still is room for a second one, but producers of a third or fourth system have very little chance of ever making a profit. A company exclusively possessing a standard has a near-monopoly. In this way, Microsoft holds 90 percent of the market in PC operating systems, Intel holds 90 percent of the market in microprocessors, and IBM holds 83 percent of the market in mainframe computers, and practically the entire non-Japanese market for mainframe operating systems.

Qwerty, the most boring standard ever

The qwerty-standard is the classical example of a standard that is impossible to replace. Samuel Sholes, an American, invented a typewriter in1868 which suffered from 2 problems: the little hammers got stuck together when he tried to type fast and the machine needed to have a special trick he could demonstrate. By using the QWERTY configuration Sholes put the most often-used keys widely spaced, so the little hammers wouldn't get in each other's way. This also made it possible to very quickly type the word 'typewriter' on the top row – a handy sales trick.

All Sholes still needed then was someone to produce the machines. He found an ally in arms producer Remmington, who was looking for new businesses after the Civil War had ended. To help sales of the typewriter, Remmington organized typing contest, which competing machines also entered into. Remmington contracted the winners. It was only a short while before typists invented touch-typing, using 10 fingers. Educational institutes quickly adopted the system, and soon everybody wanted the qwerty system: companies because their secretaries could work quickly using the system and the educational institutes wanted it because most companies used qwerty machines. Despite the fact that there was no longer a technical necessity for the QWERTY arrangement, most other manufacturers had also adopted the system by the turn of the century.

Since then a few alternatives to QWERTY have been designed. The most well known of these is the keyboard developed by the American ergonomist August Dvorak in the thirties. Using the 'Dvorak Simplified Keyboard' people learned to type twice as fast, learned to type at a rate twice higher than the QWERTY people and experienced a twenty-fold decrease in hand strain. Some Dvorak-machines were made and the configuration is optional on some apple computers, but despite all that Dvorak was no success.

After Dvorak even better keyboards have been developed, but none of them managed to push QWERTY of the throne. Qwerty had installed itself into the brains of millions of people and couldn't be erased from them.

Why hasn't that happened? Surely an ergonomical layout would have prevented lots of stress and maladies like repetitive strain injury (rsi)? And it can't be that expensive just to plug in a new keyboard into your computer?

A number of explanations can be given for the survival of the QWERTY system. Firstly, 'qwerty' can't be unlearned. Touch-typing is never unlearnt once learnt, in the same way swimming is. People who touch-type won't be quick to switch keyboard layouts.

A second explanation is that the improvements in learning speed, typing speed and hand strain apparently aren't sufficient reason for young people and typing institutes to justify switching to a new keyboard layout. Apparently typing is supposed to be difficult and hard to learn. Experts on innovation state that a new standard should be about 10 times as good as an old one in case it is to be adopted.

Furthermore, vested interests play a role. Why would educational institutes be interested in learning their students to type in a shorter time? Their income would only suffer and they would also get into trouble with companies who would have to purchase a new machine or software for use by their new secretary.

Increasing returns

The Economist recently characterized these near-monopolies as 'technopolies'. The classical

law of decreasing returns doesn't hold in the case of technological standards. According to that law, a product which's sales volume keep increasing reaches a ceiling, beyond which profit doesn't increase any further. The causes for this are increasing costs per unit sold and the launching of alternative products by competing brands. This mechanism keeps prices low and prevents excessive market shares from being reached.
On the other hand, in the case of technological standards, the situation is one of increasing returns. Sales start off only slowly, but when in the end everyone switches to standard market shares soar towards 90 percent. After this phase any potential competitor has quite a challenge to rise up to.
Economist W. Brian Arthur from Stanford University describes how a network of new companies forms around every new technological product[23]. In Microsoft's case its computer manufacturers, software developers and microchip producers and in the case of photographic films its camera manufacturers, film developing services and camera film producers.
Paul David, also an economist at Stanford University, calls the phenomenon 'system scale economies'. Users desire a product of a uniformly valid standard, so they can easily obtain compatible software and exchange information with others. Because of this high 'networking' tendency, a small advantage in the early developments of such a standard will persuade more and more people to buy the product, after which sales increase by themselves. Production costs drop, while those of the competitor rise.
Technological standards always enter the market in pairs, David asserts; a software part and a hardware part. Software anchors itself in the market, and gives the hardware part leverage to increase sales. When consumers have bought the software part, they are committed to that standard. Nobody will decide to throw away his CD collection in a hurry, or to exchange all Windows machines in the office for Apple computers.
The heavy competition amplifies the snowball effect. Manufacturers lower their prices, more people buy their new products and hence create an even larger market for the winner. Netscape and Microsoft were even giving away their products to win the battle for the Internet browser market, making it expand rapidly in the nineties.

End-game strategy

There can be only one winner, but how is the game played? The most important part is the end-game strategy. A business interested in making its technology the standard on the market, shouldn't try to pursue short-term profits. In this way, Philips gave away licenses in the sixties to everyone wishing to produce its music cassettes. Microsoft signed an agreement with IBM in the early eighties that didn't yield it much profit, but ensured that its operating system MS-DOS became the standard.
Besides the end-game strategy there are more specific explanations for success or failure. For instance, the term 'quality' of a product must be interpreted in a loose setting. Strictly, the sound quality of a CD is actually worse than that of an LP (nuances lost by discretization of the audio data), but other aspects of quality determined the outcome: the longer playing time, the smaller size, the ease of use and the fact that CDs are more robust.
Proper marketing also is essential. When it introduced the CD, Philips covered up the worse sound quality of CDs in a clever way by emphasizing all the advantages they had over classical LPs.
Entrepreneurs in the computer industry recognize an important factor responsible for the

[23] W. Brian Arthur, 1996, Increasing Returns and the New World of Business, Harvard Business Review, July-August 1996

success or failure of a new technology: the killer application. Such an application is a certain application that encounters such popularity that it persuades millions of people to buy The spreadsheet Visicalc for instance was largely responsible for the successful introduction of the PC by Apple and IBM. Killer applications also occur outside the computer industry, though: Edison invented the phonograph using wax rolls in 1877. Edison, who was more of an inventor than an entrepreneur, thought the many uses of his product formed its novelty: as a voice recorder, recording news, speeches, strange languages and oh, also maybe for music, too.

The German entrepreneur Emile Berliner foresaw music becoming the killer application. He marketed the gramophone invented by him and at the same time started a record label. Within a few years the gramophone was market leader and Edison started producing them too.

The laserdisc is a more recent example. Philips tried to market such a device for playing interactive CDs under different names three separate times, to no avail. Competitor Pioneer introduced the killer application: a laserdisc machine suitable for Karaoke, incredibly popular in Japan. Japanese sales skyrocketed, as did Pioneer's market share.

Licensing politics

Besides the end-game strategy, quality, marketing and finding a killer application, licensing policy also is an important factor determining the success of a new technology. By giving out licenses a business can quickly increase the market share of a new standard and rapidly increase the number of companies that have an interest in it.

Microsoft is well known for its extremely crafty licensing policy. When the company wanted to compete against cp/m in 1980, it developed MS-DOS at a low price for IBM's first PCs. That wasn't where it ended however, American computer journalist Robert X. Cringely writes in his book 'Accidental Empires'[24]. When the IBM-PC was already marketed, many American computer manufacturers were still producing their own brand PCs. Microsoft offered to make a different version of MS-DOS to each computer manufacturer. Whichever company would win the battle for the standard; it would have an MS-DOS operating system. As soon as the manufacturer had signed the contract, Microsoft told them that not all IBM applications would work on the system. This startled the manufacturer: a PC having no compatible software won't sell. Subsequently Microsoft divulged they happened to have a series of programs – from word processor to spreadsheet – which were easy to adapt, at a price of course. Even before the computers were on sale, Microsoft had made its profit.

Ruining the market

Within established industries, another factor is of interest: the poker game played by the big businesses dominating that industry. These businesses watch every move of the others carefully, paranoid as they are that the competition will run off with some new standard. This can make the market lock up completely, with the formation of alliances between different competitors as the only solution.

[24] Robert X. Cringely, 1991, Accidental Empires, How the Boys of Silicon Valley Make Their Millions, Battle Foreign Competition, and Still Can't Get a Date, Harper Business.

The Formula

Michael Hay and Peter Williamson[25] the following tips to entrepreneurs in a new branch:

- Don't put all your eggs in one basket. Exchange licenses in case a competing standard wins.
- Attract the renowned companies, the opinion leaders, as first customers.
- Make sure you get quick market feedback.
- Invest in production technology, necessary process technology and supporting products at the same time.
- Recognize changes in structure and composition of the competition as soon as possible.
- Think ahead towards the end game, when the branch has settled. Think of ways to maximize your profit in that stage, and don't try to make all your money in the chaotic and insecure period before then.

A non-regulated introduction of a technology can in such a situation lead to a ruining of the market. The record industry experienced this in the forties, when the American record company Columbia released the LP record (33 revs) in 1948 and competitor RCA released the single (45 revs). Both record types sounded clearer and lasted longer than the old 78-rev record, but didn't work on the same record players.
The consumer refused to make a choice and wasn't prepared to buy two separate record players either. For four years the market in the US was stuck. This 'battle of the speeds' only ended when a record player was introduced which was capable of playing both formats. At this point the record industry flourished: Adults mainly bought the more expensive LP records, youths bought the singles.
Growth of the video recorder market was hindered too by the existence of two standards. Only when it became clear that the VHS system would win over the V2000 (VCC) system by Philips, the market resumed growing at a high rate. Philips' system was technologically more advanced than the Japanese effort, but had to halt production of the V2000 system in the end. Less-than-sensible market distribution agreements, the lower standard of reliability, the bulky design of the first generation of recorders and most importantly the lack of available movies became the death-knell for the system in the end.

Tremendous interests

Large businesses realize that it is often more profitable to cooperate than it is to compete for profit among each other. Every manufacturer knows that if their standard wins, there is major profit to be made. However, to increase the chance of success technology can also be shared with competitors. Because everyone is aware of the tremendous interest at stake, this cooperation leads to a complicated game of strategy. At the introduction of the digital versatile disc (DVD), which is supposed to replace the videotape, CD and CD-ROM, it started all over again.
The electronics industry divided itself into two camps: an alliance around Philips and Sony, and another one around Toshiba and Matsushita (including Panasonic). The manufacturers initially opted for cooperation, which resulted in tough negotiations dragging on for a long time. For Sony and Philips, together holding the most important technology, this took too long. They declared they would continue pursuing their own standard past autumn. From fears for the market-ruining competition that was about to break out, the competitors decided to return to the negotiation table. Finally, they reached an agreement on a common standard.

---

[25] Michael **Hay,** Peter **Williamson, 1991**, The Strategy Handbook, London: Basil Blackwell.

Sony and Philips now rake in the biggest royalties: 2.5 percent off of each DVD player sold and 4.5 dollar cents for each DVD disc sold.
Besides the standards which form in the chaotic competitive battle on new markets and in the poker game played by established industries, there is a third category: standards enforced by the government. This generally concerns branches, which used to belong to the public sector such as television, the telephone system or the electricity grid. Businesses commit huge resources in persuading the government to adopt their system as the standard. This is why negotiations on the standard for HDTV (High Definition TeleVision) have dragged on for ages. Telecommunications companies in Japan have been busy lobbying for the new standard of the cellular phone network. In the end, the US standard was adopted.

Breaking standards

To profile themselves against their competitors, some manufacturers choose to pursue an open systems approach. Hewlett Packard for instance became successful by making products that could interface with almost any other.
Sun Microsystems is a typical breaker of standards. The company's strategy is to quickly gain market share using open systems and standards. The first computers manufactured by the company (which was founded by students) were built using existing, non-reserved parts and designs from Stanford University – SUN stands for 'Stanford University Network'. Subsequently the company designed simple software and simple operating systems, the source code of which they made publicly available. Those first programs were incredibly buggy, but everyone used them – including IBM – because they were free. This made it easy for Sun to sell the hardware supporting the programs. This wasn't the end of the story, however. Sun then encouraged other manufacturers to clone the Sun workstations, so a standard could be formed quickly. The way Sun would make its profit was by continually releasing hardware that was that bit better than that of the competition. The programming language Java, widely used on the Internet, even appears to 'open up' the competitors' systems.
The American computer company Cisco isn't protected by a standard at all. The company makes Internet routers, switchers that reroute packets of data towards their destinations over the Internet, and has a market share of about 85 percent. Users can, in principle, make the switch to one of the competitors' machines in a matter of hours. The only way Cisco can survive is by simply being better than the competition. If the company misses a new innovation, it has no choice but to merge with the competitor who has implemented it. Cisco has spent billions of dollars on such practices. The laws pertaining to technological standards don't apply here anymore – only the law of the jungle: eat or get eaten.

The formation of trajectories: positive feedback[26]

The scientific discipline that concerns itself with the question, which choices companies make when in uncertain circumstances, is evolutionary economy. According to evolutionary economy certain 'technological trajectories' or 'paths' form over time which companies can get stuck into. The technologies in which companies or economies get stuck don't always

[26] fragment from paper 'Milieustrategieën en positieve feedback: kunststofverpakkingsafval als illustratie' by Caroline van Leenders and Paulien de Jong, UvA, 1996.

have to be those technologies that are most efficient for their users[27]. Arthur states, that the technology which wins the competition for becoming the standard, i.e. gaining a major market share, doesn't have to be the best option for users in the long term. An example of users being caught up in an inefficient technology is the VHS video system. VHS 'won' over competitors Betamax and V2000, despite the fact that it was neither the cheapest nor the technically superior system.

According to Arthur, the formation of technological paths, which can be inefficient, is a consequence of the fact that the phenomenon 'increasing returns with increasing market penetration' occurs. This means that the more a technology is adopted, the more it improves and the more attractive the technology becomes for further development. A situation in which a technology has an advantage in adoption and this advantage is self-reinforcing is also called a 'positive feedback' situation.

Positive feedback: The situation in which an advantage in market share of a technology (versus competing technologies) reinforces itself (leading to market dominance).
Trajectory dependence: The dominance of one technological alternative over competing alternatives is dependent on the development trajectory of all alternatives (and not just dependent on the price/performance ratio of each separate alternative).

Arthur mentions six factors causing positive feedback, to be specific:

a. Expectations

The development of a certain technology can be influenced and accelerated by the expectations people hold as to the success of the technology. Expectations contain a certain image where a future situation is sketched, connections are made and roles are described. Based upon these expectations new actions are undertaken.

b. Familiarity

When a technology is better known and better understood, it has an increased chance of being adopted. Arthur also describes this factor as 'increasing returns by information'.

c. Network characteristics

Positive feedback shows up more strongly with technologies possessing network characteristics. It is advantageous for a technology to be associated with a network of users, because this increases availability and the number of product varieties. Again, a good example here is the VHS video system. To be able to function, this technology needs a network consisting of video rental stores stocked with VHS tapes. The more users are present, the better the possibility is for users to profit from VHS-recorded products.

d. Technological connectivity

Feedback processes are stimulated by the occurrence of 'technological connectivity'. Rosenberg already posed in 1979 that innovations depend on the existence of complementary technologies. Often, a number of other sub-technologies and products get absorbed into the infrastructure of a growing technology. This gives it an advantage over technologies, which would need a partial demolition of that infrastructure to function. An important study

[27] In neoclassical economy it is assumed that the 'best' technology would conquer the market. In Arthur's model the process is trajectory-dependent, which makes it impossible to predict which technology will conquer the market.A consequence of this is the fact that an inefficient technology can actually win the race.

regarding feedback processes caused by technological connectivity is David's research into the QWERTY keyboard (which's name refers to the first six keys on the top row of the keyboard)[28]. Summarized: a technology, which fits into the system of already existing technologies has a relatively better chance to develop than a technology, which lacks those connections.

e. Economies of scale

When an increasing volume of products is produced while the costs per unit production don't increase linearly with it, the price of a product is lowered. This means that a technology can become more economical when it is applied on a larger scale.

f. Learning processes

Positive feedback during the development of a technology can finally take the form of a learning process, because a technology can be improved more quickly when more is learned during its use. Arthur states that, when more is learnt about a technology, this technology gains an advantage in application. So when a company learns a lot about using a specific technology but learns little about another, this last technology has less chance of being adopted in the future.

For the successful innovation of environmental innovations mainly 'interactive learning' is of importance. This is a specific form of learning and is also called 'learning by interacting'. This kind of learning occurs when contact exists between different stakeholders in the development process.

[28] In this study, David asks himself why this specific keyboard has won from the competing alternatives. According to David, one of the factors of importance here is that of technological connectivity, in this case compatibility with the system.

Michael Porter, The Competitive Advantage of Nations[29]

Diamond model
The Diamond model of Michael Porter for the Competitive Advantage of Nations offers a model that can help understand the competitive position of a nation in global competition. This model can also be used for other major geographic regions.

Traditionally, economic theory mentions the following factors for comparative advantage for regions or countries:

A. Land
B. Location
C. Natural resources (minerals, energy)
D. Labor, and
E. Local population size.

Because these factor endowments can hardly be influenced, this fits in a rather passive (inherited) view towards national economic opportunity.

Porter says sustained industrial growth has hardly ever been built on above mentioned basic inherited factors. Abundance of such factors may actually undermine competitive advantage! He introduced a concept of "clusters," or groups of interconnected firms, suppliers, related industries, and institutions that arise in particular locations.

As a rule Competitive Advantage of nations has been the outcome of 4 interlinked advanced factors and activities in and between companies in these clusters. These can be influenced in a pro-active way by government.

These interlinked advanced factors for Competitive Advantage for countries or regions in Porters Diamond framework are:

1. Firm Strategy, Structure and Rivalry (The world is dominated by dynamic conditions, and it is direct competition that impels firms to work for increases in productivity and innovation)
2. Demand Conditions (The more demanding the customers in an economy, the greater the pressure facing firms to constantly improve their competitiveness via innovative products, through high quality, etc)
3. Related Supporting Industries (Spatial proximity of upstream or downstream industries facilitates the exchange of information and promotes a continuous exchange of ideas and innovations)
4. Factor Conditions (Contrary to conventional wisdom, Porter argues that the "key" factors of production (or specialized factors) are created, not inherited. Specialized factors of production are skilled labor, capital and infrastructure. "Non-key" factors or general use factors, such as unskilled labor and raw materials, can be obtained by any company and, hence, do not generate sustained competitive advantage. However, specialized factors involve heavy, sustained investment. They are more difficult to duplicate. This leads to a competitive advantage, because if other firms cannot easily duplicate these factors, they are valuable).

[29] http://www.valuebasedmanagement.net/methods_porter_diamond_model.html, February 16th, 2005

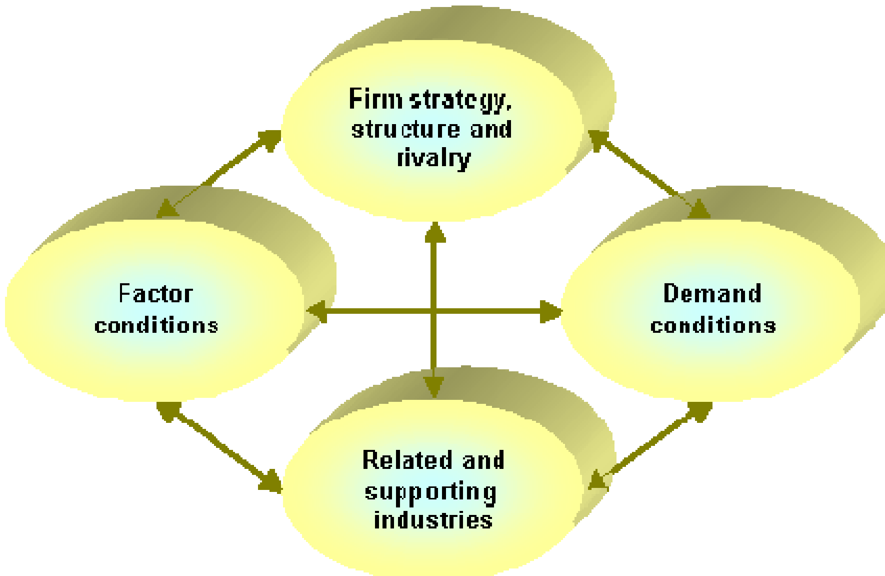


Figure 1 **Porters' diamond (http://www.themanager.org/Models/diamond.htm, February 16th, 2005)**

The role of government in Porter's Diamond Model is "acting as a catalyst and challenger; it is to encourage - or even push - companies to raise their aspirations and move to higher levels of competitive performance …". They must encourage companies to raise their performance, stimulate early demand for advanced products, focus on specialized factor creation and to stimulate local rivalry by limiting direct cooperation and enforcing anti-trust regulations.

Porter introduced this model in his book: The Competitive Advantage of Nations, after having done research in ten leading trading nations. The book was the first theory of competitiveness based on the causes of the productivity with which companies compete instead of traditional comparative advantages such as natural resources and pools of labor. This book is considered required reading for government economic strategists and is also highly recommended for corporate strategist taking an interest in the macro-economic environment of corporations.

Methods and Tools

In this chapter several methods will be treated with which can be attempted to chart future technological developments and their consequences. With that the matter of how much can be predicted in a certain situation will also be discussed. Finally the 'control dilemma' will be treated: in the phase when the influencing of technological development is still possible, we aren't sufficiently (or not at all) aware of the social consequences. When we finally do know these consequences, the technology has set beyond the possibility of adjustment.

Predictability

How can we make sure that predictions (or in the more general sense: statements being predictive in nature) are more dependable than making a bet in the casino? Predictions have been made for millennia in the past. Sometimes they were religious in nature. Their pretence to truth usually wasn't irrational, but extra rational, which means their claim to truth exceeded the boundaries of rationality (oracles, revelations, visions etc.). These predictions invariably had the character of 'fate' in the sense that they would certainly come true and couldn't be influenced by any actors. Although belief in the 'customization' of society has drastically reduced in the past twenty years, no valid reason exists to reduce it to 'zero'. However, the kinds of predictions involving fate are unacceptable to us for two reasons:
The predictions are to have a rational foundation, meaning a plausible insight must be given into causal relations, including possible feedback loops, on which the prediction is based.
Predictions are to have a conditional character, meaning they should indicate the conditions under which the consequences (possibly) will occur and hence also possibilities to influence the course of future events.
Predictions must be useable. Useable in this context means: answering a question with a relevant answer, which is credible and communicable and is given at the right time.

Methods

There is no such thing as a perfect foresight method because the relations on which the prediction is to be based often hold an empirical or inductive character (meaning extrapolations from observed relations in the past). It is hence principally uncertain whether these relations are to remain valid in the future. It can even be a case of a clearly measured relation between two variables only turns out to be a coincidence of environmental factors. Therefore a study should preferably adhere to a 'theoretical framework', which indicates under what boundary conditions the empirical relations used are valid. There is no method that leads with utmost certainty to crystal-clear predictions; only methods that lead towards verifiable conclusions.
In Foresight mistakes often occur. Common mistakes are:
research hypotheses aren't stated explicitly and aren't made credible. This can for instance lead to circular reasoning; a conclusion has been implicitly stated in the research hypothesis.
observations are too readily considered undebatable. Every observation depends on the conceptual framework of the observer and can always be debated in principle. Researchers should acquaint themselves with this, especially when it concerns the observation of social phenomena. This holds quite often for the judgments of experts, especially where they are judging technologies that threaten 'their own field'.
In the past a number of methods and resources were used in foresight studies. None of these

methods is adequate to base an entire study upon; combinations were practically always used and required the addition of literature studies and interviews with experts. These last two methods are so commonplace that they won't be discussed extensively in this section. However, some tips are important here:
Never blindly act based upon the opinion of one single expert. Experts have interests of their own which makes them have a specific vision. Also, they sometimes give advice on subjects that they aren't experts in at all.
The memory of people interviewed is often rather limited. Sometimes they can proclaim fallacies with the utmost certainty. Therefore, always keep a critical eye on statements made in interviews.
Ask for specialist advice on the literature study. The number of databases is so large that it would be impossible to find all the relevant data within a reasonable time span.

Monitoring

Monitoring: the continuous observation of all public information on a technological area consisting of channels like publications, trade fairs, lectures, annual business reports, patents etc.
In monitoring an eye is continuously kept open for all signals, which may play a role in developments under interest. This includes for instance magazines, subject literature, media coverage, business periodicals, congress reports, advertisements, annuals, databases etc. Talks with experts can also be quite useful. Specifically 'gatekeepers', meaning people who receive information from lots of different angles, are of importance. Monitoring a technology in this way is very intensive. A special way of monitoring is the bibliometrical analysis. Here databases are screened using specially designed computer programs. Technological trends can appear from patent files. Also shifts between countries can be observed, as well as a shifting of the focal point within disciplines. Forms of cooperation and the forming of new issues can be analyzed. These methods form a useful tool for detecting trends, but can hardly say anything about their meaning.[30]

The Delphi method

In a Delphi a number of experts are interviewed in writing on the development of a specific technology. The experts subsequently are presented with the arguments and estimates of the other participants (anonymously). Afterwards they are again interviewed, especially regarding the items they showed disagreements on. Usually a consensus is reached after about 3 rounds.
The Delphi method is a method using which the expectations of experts on the development of a certain technology can be charted. With a Delphi a number of experts are required to answer a series of questions on the development of a specific technology in writing. The response should contain a judgment involving a timescale and an estimate as to the probability of future developments. A project leader gathers all responses and poses a new, more specific series of questions to the people interviewed. In this new round the matters on which consensus hasn't been reached are the most important. The participants get to review the arguments and estimations of the other participants anonymously. In an iterative process it is attempted to reach a consensus among the experts. Using this a more solid foundation is gained for a vision of the future. These rounds force the experts into refining and supporting

30 Caroline van Leenders, 1992, Milieugerichte technologieverkenningen inventarisatie en selectie van verkenningsmethoden, TNO-STB.

their argumentation.
No face-to-face discussions take place, but the matter is treated in writing to prevent verbal trickery from being used and to keep status from playing a role. In practice usually no further convergence of opinions occurred after about 4 rounds. Sometimes a forum discussion was organized between participants because of waning response. The Delphi method is particularly suited to foresights in which the opinion of experts plays a major role.
Methodologically a Delphi comes with its own pitfalls. How can be prevented that a bias occurs because of lack of response from a particular group, for instance industrial experts? How can it be prevented that experts try to contact each other outside of the study, in worst cases even to discuss on the answers they will both give in the interviews, to consolidate their own interests (for instance more funding or prestige for their work, or the sweeping under the rug of problems which can lead to unwanted government interference)? Moreover a Delphi takes a lot of time, both for the analysts and for the people interviewed.
It seems questionable to involve Delphi participants from different interest groups. In a conflict of interests rational arguments don't play the leading role. Is consensus still attainable in that case? Are the participants still using the same standards?

Cross Impact Assessment

A variation on Delphi is the Cross Impact Assessment method, which was first used by the American aluminum company Kaiser, mainly in decisions on the introduction of new products. For the company 60 developments were identified which might happen at some stage. In constructing these, various experts were involved to chart risks in terms of technology, competitor and client behavior, etc. They weren't only asked to estimate the chance on such a possibility happening, but also on what the chance on its occurrence would be if some other event had happened before. The result was a matrix of chances. This matrix could be calculated using a Monte Carlo procedure, which made it easier to assess the risks following from a decision in a certain complex situation.
Using the Cross Impact Method risk-avoidance measures could also be devised. This method appears to be suitable mainly for strategic studies because it isn't so much about exploring a new area as to reduce the complexity of a problem. Cross Impact Assessment can in principle also be done using other elements besides the opinions of experts.

Example of a limited events matrix regarding the introduction of the fax machine on the market

Probability of event becomes:	If this event happens:			
	Increasing taxes/costs	Negative legislation	Replacement technology	Market saturation
Increasing costs/taxes on transmission (0,70)*	1,00 P(1|1)	0,40 P(1|2)	0,70 P(1|3)	0,51 P(1|4)
Negative legislation (0,40)*	0,20 P(2|1)	1,00 P(2|2)	0,38 P(2|3)	0,31 P(2|4)
Development of replacement technology (0,60)*	0,90 P(3|1)	0,72 P(3|2)	1,00 P(3|3)	0,33 P(3|4)

Market saturation (0,45)*	0,33 P(4|1)	0,35 P(4|2)	0,05 P(4|3)	1,00 P(4|4)

* Initial marginal (ceteris paribus) probability.

Socio-technical maps

The social-technical map shows:

- The state of development of a technology,
- The dynamics in development of this technology,
- The different stakeholders involved in this technology,
- The views and interests the stakeholders have regarding this technology.

To not only be able to predict technology but also give information on social effects, more tools are needed. The views of groups and organizations, which are relevant, form the basis for social-technical maps for specific technological developments. Below a list is stated using which such a map can be formed.

0. Bounding of the technical system (will it be a map of the car, the engine or for instance the electrical engine) and time frame.
1. Construction of a crude tree showing hierarchy of technical alternatives and mechanisms, which determine the selection between them, plotted against time, meaning which alternatives are being worked on and what choices are made in the process. Attention is also paid to alternatives that didn't make the grade, but can be developed on separate tracks in other developments.
2. Characterizing the alternatives according to contents (cognitive) and origins (social). Which stakeholders are trying to get which items onto the agenda? In the characterization of contents attention is also paid to expectations, links between alternative technologies (for instance coupling to base technologies) and any missing knowledge. In the characterization of origins also focused upon are relationships between stakeholders generating alternatives and the identity of the relevant stakeholders.
3. Does trajectory (subsequent technologies are all based on the same basic design/knowledge base, like C-MOS technology, see example) formation occur? Trajectory formation is often linked to expectations regarding technological progress, which results in disregard for alternative technologies. Trajectory forming occurs together with entrenchment (see SCOT model, the 'nesting' of technologies in our society). This means that various different actors (legislators, research organizations, or even consumers) get involved in the project and adapt themselves towards this development.
4. What are the (environmental) effects of the different alternatives? When has who acknowledged these effects, and how are they taken into account?
5. Are there critical episodes visible in the technological development? How can those fractures be characterized in both a cognitive and social way? Through which process did the fractures occur? What roles do the different stakeholders play in this? Was it a matter of anticipation by technologists, or external pressure? How was the social environment affected by the technology? In what way were the technologists put under pressure by their social environment? Were there specific actors linking the technology and its social environment?
6. In the periods where no fractures were present (developing episodes), were there attempts to bring about fractures? By whom, and why did they fail? What difference is there with the critical episodes?

Checklist

The making of a social-technical map often takes a rather large amount of time. The technology has to be monitored intensely and a large number of interviews are needed to chart the positions of the different stakeholders. Often a quicker check is wished for, to see if there should be cause for concern and to come up with subjects for further research. The checklist presented here consists of three parts (unequal in weight):

- research and development work
- product
- production process

'Product' in this case means that which is the intended final result (and usually is sold), 'production process' means the steps taken to produce this product and 'technology' signifies the product, production process and knowledge incorporated into these.

A: How acceptable is the research and development work that accompanies the formation of the new product and its production process?

1) Do counteracting social forces exist against the methods used in the research and/or development work or against the collection and storage of certain data?

2) Is the research and development work scientifically interesting or does the development of this technology provide a special contribution to (a) technical and/or scientific discipline(s)?

3) Can it be foreseen that investing in the development of a technology at this moment could prevent a better alternative from being developed in the future?

A1) These points can be considered in this case:
Safety of the research for researchers and people living in the neighborhood regarding for instance the release of poison, manipulated organisms, radiation etc.;
Abuse of test animals or test subjects (people);
Possible forms of abuse of research data for socially controversial ends such as weapons of mass destruction, gaining access to private information, race-based discrimination etc.;
Mistrust towards researchers, which are carrying out the research (a so-called Faustus, Jekyll, Strangelove or Buikenhuisen complex).

A2) Disconnected from the eventual results a research or development project can sometimes provide a powerful stimulus to other technologies and disciplines. Many examples are known. For instance, the construction of the Oosterscheldt storm surge barrier in the Netherlands yielded new knowledge and technology, which could be used for other goals. This effect was also to be seen with the SDI project. These effects are usually called 'spin-off effects' when coming from military technology.

A3) This problem occurs quite often: investing in an improved waste-incineration process that prevents the forming of dioxins makes developing a more environmentally friendly alternative to PVC later on less attractive. Investing in a more efficient petrol engine means that the development of the electrical car will suffer as a result.

B: How acceptable is the new product in itself?

Ethically:

1) Are any social values connected to the product in itself, or the product that it replaces?

2) Is the product considered unacceptable in the ethical system of specific religious or cultural factions?

B1) People don't judge products based solely on their own interests; they also rate the product in terms of their social opinions. Products can expect to receive a certain amount of positive appreciation when they can be coupled to social developments/changes, which are considered to be positive. In this way many people in the neighborhood of Medemblik in the Netherlands held a positive attitude towards the construction of a new 1 MW windmill because they expected a positive environmental contribution from it.
A similar coupling is suggested for the appreciation of 'atomic electricity'. Here, two conflicting social values were in play: economic growth and safety/health. Analyses show that many of the differences surrounding the use of nuclear energy corresponded to views held with respect to these social values.
A negative appreciation could for instance exist for products which can be related to negatively valued social phenomena such as:
Unemployment (this was an issue when the computer became hugely popular);
Forms of opulence and waste (probably the reason for the failure of the electric toothbrush in Holland in 1973 by Philips, and their subsequent decision not to introduce an electric corkscrew);
Animal cruelty (fur coats);
Usurping of traditional culture (replacement of windmills by engine-driven pumps in the nineteen-twenties and thirties).

B2) This particularly pertains to protests made by religious factions, for instance against certain kinds of foodstuffs, preparation methods, contraceptives, medical treatments and animal products. Such protests don't necessarily have an influence on acceptance by the majority of the population (for instance the Roman-catholic protests against the birth control pill).

Social acceptance

3) Will the costs of the new product be likely to attract criticism?

4) What effects on the environment will the product have? Which changes in behavior will the product cause, and what environmental effects result from that?

5) What are the risks the use of the product entails, both for the user and others?

6) Does the use of the product clash with habitual behavioral patterns of large groups of people?

7) Are there financial or psychological barriers which hinder acceptance of the product?

B3) The cost is partly a technical/economical boundary condition. For some products it can be economically acceptable to have a higher selling price than the one they currently have (for a product of equal value) because for instance legal measures can be taken to limit competition, or because competition is hardly possible at all. However, not every price that can be set market-wise (especially concerning products that can be monopolized) is also socially acceptable. Often social values play a role in this, such as in the debates on the price of university and school fees.

B4) Readily visible environmental effects stem from the use of energy by the product (and the change in it compared to that of an existing product). However, change in behavior is also important: consider an increase/decrease in car use, etc.

B5) Items here include damage to health, economical damage, psychological damage as a result of the improper functioning of the product. In this case the appreciation of risks differs: Voluntarily taken risks are much more acceptable to people than risks imposed on them by others. See for an example the famous study by Nader (1965) about the lack of safety in cars.

B6) Here, the change of commonly accepted behavior is the issue, for instance in the case of the introduction of bio-bins and glass containers. In both these cases this went fairly smoothly because people were motivated for a higher value (concern for the environment). The Unilever product 'Dentabs' failed however, because it infringed upon the deep-routed habit of teeth brushing which was instilled in the Dutch population by way of commercials. Brown milk bottles (which were less transparent to light) failed because consumers couldn't see whether the bottles were clean anymore. Also, a new product can have 'hidden' deficiencies that only show up later, after prolonged use (e.g. many people find it uncomfortable to read long texts from a microfilm or a screen). Common habits are hard to change, even if a change would be beneficial.

B7) New products can sometimes pose advantages but nevertheless have difficulty being accepted because the customer's inhibition towards using it is too great. This inhibition can consist of a course required to get acquainted with use of the product (typing courses and new software spring to mind), the necessity to purchase special equipment before the product can be used, or psychological barriers. This was in a sense the issue with new aramid-reinforced car tires, which had to compete against the 'Stand on Steel' campaign by Michelin when they were introduced in the seventies.

Secondary social effects

8) Does the product permit new (economic or otherwise) activities? How should these activities be judged?

9) Does the product threaten existing activities, which hold a certain social or cultural value?

10) Does the product influence the social structure? (private life, local community, cultural region)

11) Does the product have any other (possible) uses than the one primarily intended for it?

B8) New products can sometimes lead to a host of new possibilities, which promote the acceptance of the product. Copiers not only replaced carbon paper, but also led towards a vast increase in the amount of copying done – something IBM hadn't anticipated. In this way it is to be expected that new and improved products not only replace existing ones but also introduce increased functionality and with it a broader acceptance.

B9) A new product can lead to a decrease in demand for other products. This can cause the market for those products to become too cramped. Sometimes these products are considered too valuable to disappear: think of reduced theater visits as a consequence of the use of television, or a reduction in the use of public transport as a consequence of increased car use. The community considered these products (theaters, public transport) so important that they were often subsidized.

B10) Many products exert an influence on the way people live together and communicate. Where once the latest rumors were exchanged at the pump in the village square, there now is the local cable TV network. This means for instance the introduction of a partial news monopoly and the loss of a part of the local social structure. The television also meant an attack on family life. Also consider 'computer widows', scientists who daily email with their colleagues and commute, etc.
Often these drawbacks of a new technology pose no problem towards its acceptation: the decision to accept is often individual (as are the advantages); the drawbacks (the dereliction of social communities) are usually collective.

B11) Especially base products often have many possibilities of use. Polyethylene for instance was originally developed as an insulator for submarine cables. This later only formed a fraction of its total use.

C: How acceptable is the production of the new product?

In itself:

1) Are any ethical standards and/or values threatened in production?

2) Are the working conditions in the production process acceptable?

C1) Consider for instance (miss-) use of animals or people, and religiously inspired protests against the use of holy ground and violation of the rest on religious celebration days. Also to be considered are forms of resistance against production processes that mainly stem from negative associations and emotional responses. This plays a part in for instance food conservation processes in which radiation plays a role. Next to very concrete arguments on the effect radiation has on food, ethical values on the value of life, creation and future generations also play a role in this case. Ethical standards and values also crop up in the acceptance of food, which has been produced by means of genetically modified organisms. Cheese, which is produced using chymosine obtained through genetic modification, is identical to cheese produced using rennet obtained from calves' stomachs. Yet the 'Gist-Brocades' company is unable to sell chymosine because (mainly German) consumers don't wish to eat cheese produced in this way.

C2) Here attention has to be paid both to the physical and the psychological work environment. Workload, safety, stress, exposure to hazardous substances, continuous shifts, possibilities for employees to be involved in the organization.

local environment:

3) Which are the physical effects of the production facility on the environment?

4) Which are the expected (primary and secondary) effects of the production on employment? What level of schooling is required for personnel?

5) What other consequences does the product have for the local environment?

6) What are the social implications of the production for the local community?

7) Does a potent breeding ground for local activism exist?

8) Can choosing a suitable location drastically reduce negative effects?

C3) This case concerns itself with the environmental impact which doesn't play a big role in the overall environmental picture, but can lead to severe problems locally, such as the emission of polluting substances into the air (it's mainly the 'stench' factor in this case), surface water or soil, a local garbage disposal problem, noise, use of precious space, 'horizon pollution', interference of electromagnetic signals, depletion of ground water, lack of safety (think of sabotage too), disruption of animal life.

C4) There are both employment effects directly related to production (both for new as well as disappearing employment) as well as indirectly related effects. Examples of indirect effects include: farmers losing their land, the building contractor building houses for his workers, a chips stand in front of the factory gates, etc. New production activities can also draw other industries into the area. The level of education of the personnel needed often is very important regarding the possibilities of local employment and the migrations caused because of it.

C5) Here secondary effects like transport and traffic risks as a result of supplies and distribution on the location of production come into play, but also possible cooperative use of the infrastructure built for the production facility. This cooperative use can lead to an improvement of the traffic situation or to better public facilities. Also waste products (like waste heat) are sometimes useable by the local community. These effects can't be very clearly categorized at an early stage, normally.

C6) In this case can be thought about the consequences of migration and industrialization on the local culture. Does the local community have an 'open' culture? Here the existence of a local industrial or trading tradition to which the new activity can add is important. Also, the effects of migration on the local housing market can be of importance. Other local consequences can be found in the area of public amenities and local taxes.

C7) Local acceptance of nuclear power plants in the American situation was linked to certain social-economic and political characteristics of the area. Particularly the already present environmental lobbying activity appeared to be directly linked to the strength of opposition against nuclear power plants and a higher living standard (high average income, few welfare-supported) appeared to be inversely proportional to it.

C8) Some of the negative effects of production will be negligible when the production takes place within for instance a large industrial area. A major urban agglomeration also offers different possibilities from a rural area. On the other hand some rural areas may offer the best environment for new production.

Society:

9) Which (either existing or planned) economic activities are threatened by production?

10) Is the existing balance of power influenced by new production? Consider the following relations:

- between employees (or unions) and employers;
- between different producers;
- between producers, clients and suppliers;
- between government and industry branch;
- between different governmental institutions.

11) What does new technology mean for the development of third-world countries? Are relations between global trade blocks influenced by production?

C9) Items to consider here are unemployment, destruction of capital (both private and public) in this and other sectors. Reduction of employment opportunities and/or investments can lead to powerful protests when it concerns groups that are well organized. When current employment opportunities can be maintained and the destruction of capital can be avoided, these protests can be soothed.

C10) Technologies sometimes can radically change the balance of social power. This is related to C4, also. The position of employees or the union can be undermined when tightly organized professions become obsolete (such was feared in the graphics sector), the government gains power over the civilian (and often also over lower-level governments) by the linking of databases, and a company can sometimes reinforce its position regarding its suppliers or clients. This can lead to resistance within the groups losing their position of power. Infamous is the resistance of British unions against some technological changes (for instance in the mining industry).

C11) New technology can spell the death-knell for the development of regions in the Third World. A more economical use of resources born of environmental considerations (or a replacement by other resources such as was the case with phosphates in detergents) can often lead to a dramatic reduction in resource exports from the Third World. New technology can also have consequences in trade politics.

Scenarios

In a scenario a possible future is portrayed for an organization, nation, discipline or

technology. The scenario should be credible and tantalizing as a possible development and should therefore be consistent and sufficiently detailed. The goal of a scenario usually isn't to make plans for future events, but to stimulate creative thinking about the future.
Scenarios are often mentioned as one method. Constructing a scenario however can hardly be called a method of forecasting because the future data that scenarios present are produced by other means. Scenarios would be better described as a way of presenting technology forecasting than as a method. In a scenario it is attempted to plot the choices, or alternative events expected and to translate the consequences of a choice or event to later choices or events (a choice often involves the elimination of a later possibility). There are always multiple scenarios possible. The determination of a limited number of scenarios which are more likely to occur than others is to be done using a different method, as should the analysis of which choices would eliminate each other. The most important goal of scenarios however is not to predict, but to 'wake people up' and make them aware of possible changes:
During stable times, the mental model of a successful decision maker and unfolding reality match... In times of rapid change and increased complexity, however, the manager's mental model becomes a dangerously mixed bag: rich detail and understanding can coexist with dubious assumptions and illusory projections[31]
Trend scenarios show developments that are in line with our current ideas. They are also called 'surprise-free scenarios' because they do not incorporate any sudden and unexpected events. The scenarios are normally shown as surrounding a most probable scenario (which often represents 'business as usual'). Scenarios often make complex problems clearer to policymakers. Shell often works with three distinct scenarios: 'business as it used to be', 'frustration and conflict', and 'realism and restraint'[32].
Framework-determining scenarios are meant to show possibilities for reaching an ultimate situation that is not an extrapolation of current developments[33] (for instance the trend-breaking scenario of traffic and transport of 1988[34]).
Beside these methods a great number of other tools and methods exist to make more quantitative predictions.

THE SCENARIO APPROACH - SHELL EXPERIENCE

It is hoped that the above section provides adequate evidence that much of methodology for decision and strategies has evolved in dangerous directions, perhaps into an evolutionary cul-de-sac. Shell experience since the early 1970s might perhaps provide an alternative route.
Although the full danger of using forecasts was not apparent back in the early/mid 1970s, the criticisms levied against forecasts were sufficiently persuasive to make us examine the whole concept of forecasting, with the result as described. lt was then that work started on the development of "scenarios", at the time a somewhat less over-used word than today. The definition of a scenario - as used in Shell - is a description of one of a number of possible

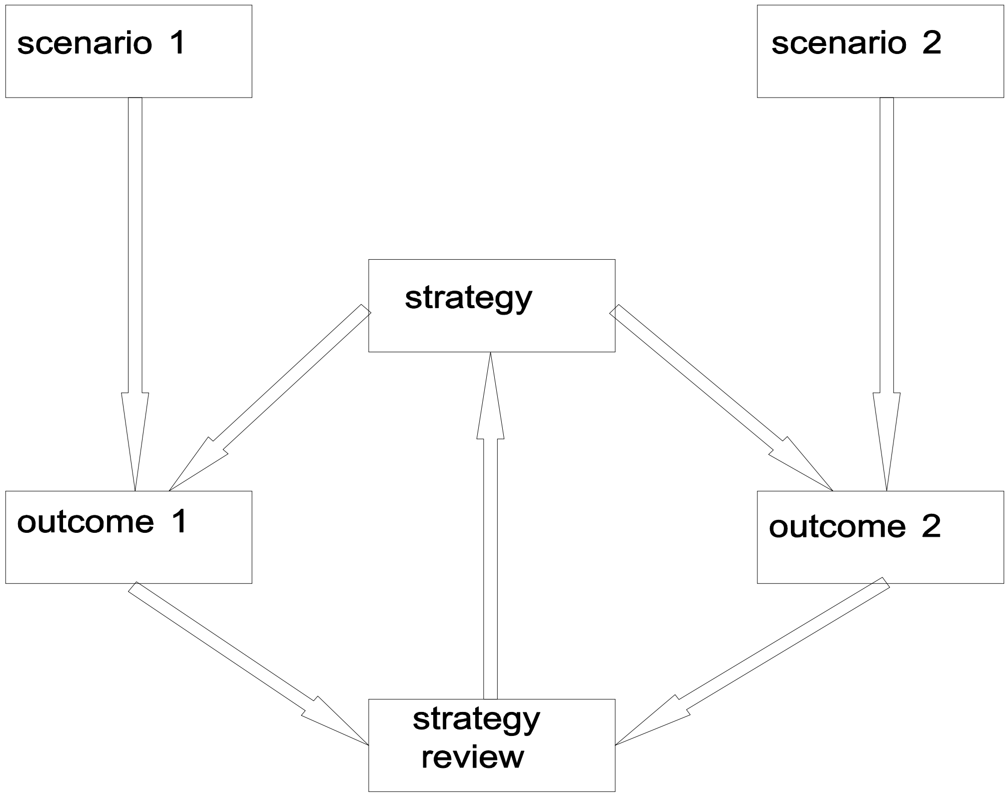
31 Pierre Wack, 1985, Scenarios: Uncharted Waters Ahead, Harvard Business Review, September October.

32 P.Rademaker, 1981, "Toekomstverkenning in het bedrijfsleven". In: Van Doorn/Van Vught, pp. 170-189.

33 Joseph van Doorn, Frans van Vught, 1978, Forecasting, methoden en technieken van toekomstonderzoek, Van Gorcum, Assen/Amsterdam.

34 Th.J.H.Schoemaker, H.C.Van Evert, M.G.Van den Heuvel, 1988, Trendbreukscenario vervoer en verkeer, TU Delft, P.M.Peeters, 1988, Schoon op weg, naar een trendbreuk in het personenverkeer, Amsterdam, Milieudefensie.

futures in which the assumptions about social, political, economic and technological developments are all consistent with each other. It is not just one of a number of forecasts.
When attempting to use such scenarios in the development of strategies or for taking decisions, it was initially found that people tended to use them as if they were a series of forecasts. So with three scenarios they would use the middle one as a "base case" and then use data from the outer two scenarios for sensitivity testing. It was then felt that perhaps one had better use an even number of scenarios, but that was no better. Four or more scenarios could not be comprehended and therefore caused confusion; when two were used, users made their own base case somewhere in the middle. Such experience showed that the whole process of decision-making was "hooked" on the use of forecasts and availability of scenarios did not change that.
But help came from a different direction. It was found that the more interesting, the richer the description of the scenarios, the greater was the appetite for discussing them. Give people a set of numbers and they can rarely comment. Give them a set of numbers plus a description of how the various forces say in OPEC might adjust themselves and so affect the oil price, and follow that by a description of a different scenario where OPEC might behave differently-and there is a good chance of a debate.
Managers, and especially senior managers, having participated in such debates, found that they achieved greater understanding of their environment and found it far easier to accept uncertainty as a normal fact of life, not just as an exception which will go away. So they found the experience interesting and rewarding.
It was also found that debate of the "unthinkable" became possible. Previously, future dangers within the environment tended to be ignored as long as people felt that the probability of them coming about was relatively small. lt was somehow felt that taking such factors seriously made them more probable (whilst one could plausibly argue the opposite).
Thus, as people got used to debating their strategies in the context of different futures, it was found normal to debate a strategy in the context of a disturbing contingency or discontinuity, even though their probability might be deemed to be low.
An even more important development arose from the fact that managers became sufficiently interested to wish to get involved in the development of scenarios. Indeed, in Shell U.K Ltd., managers now insist that scenarios must not only be interesting but must be relevant to the key issues they think about, as well as sufficiency testing for the strategies under discussion. So now the scenario writer has to involve his client to identify key issues. This is done via debate and it is astonishing how many times the issue first in the mind of the manager ends up not being his real basic concern. Once there is clarity about the issues it is then possible to "focus" the scenarios on the decisions to be taken.
As one cannot foresee whether any one scenario is likely to come about, it is pointless to develop a different strategy for each scenario. Instead, the way scenarios are being used is shown in the Exhibit.



EXHIBIT

It is at that stage that consideration of strategic options and assessment of risks have a place. But there is great reluctance to provide managers with probabilistic estimates of the future. By the mere fact of assigning probabilities an attempt is once again being made to predict the future. Risk assessment must therefore be done by the decision-makers, and if probabilities have to be assigned, this should be done by them. Debate between the scenario writers and the managers is a very useful tool to improve the managers' judgement. But the advisers must not usurp the decision-maker's role.

It is often said that the scenario approach, by leading to more robust decisions, results in risk minimisation and conservative management.

In practice the opposite tends to be the case. The more the decision-makers feel that they understand their environment, understand the dangers, they also perceive potential opportunities, the more are they able to take riskier decisions.

Experience with scenario planning has thus shown that it is not the availability of scenarios that is the key. It is their use and the collaboration between managers and advisers in their development which has caused a change in the philosophy of decision-making in Shell. This philosophy accepts that uncertainty cannot be planned away; what matters is cultivation of better judgement. It has speeded up the process of changing managers into businessmen/women.

Of course, many factors other than just the external environment come into the development of strategies - competition, technology, finance, to mention a few. Once there is an appetite for debate, is stimulated in Shell by the use of scenarios, all these other factors naturally take their place in that debate.

So after ten years' experience one might redefine scenarios as used in Shell. They are seen as perceptions of alternative future environments against which decisions are played out and as such, they have become part of a cohesive language for the whole organization.
In the author's view, it is through improving the level of debate, and the quality of thinking throughout the management structure of most Shell operating companies, that the quality of decisions taken has improved. This does not mean that mistakes have been eliminated. The only way to achieve that is through death. But the fact that Shell has survived the turmoil of the 1970s, perhaps better than many investment analysts would have considered possible, gives one some confidence that there is substance to this view.

CONCLUSIONS.

The paper has attempted to show the following:

1. *Whatever the numerical back-up, decisions require judgement-and good decisions require good judgement.*
2. *Prediction of many factors of importance to strategic decisions cannot be made reliable.*
3. *Decision-makers have to accept uncertainty as a fact of life which cannot be planned away.*
4. *They have to re-establish their power in decision-making, instead of abdicating much of this power to their advisers.*
5. *They must once again start using their judgement, just as the entrepreneurs and statesman of old. But judgement must be based on knowledge and understanding. So the manager has to cultivate his judgement.*

What then is needed is an approach which provides decision-makers - at whatever level - with:

1) *an understanding of the forces driving the system in which they operate;*
2) *a means of acknowledging genuine uncertainties and indicating what factors are already pre-determined;*
3) *ways of directing attention towards potential discontinuities so that they can prepare against them and evaluate their effect;*
4) *a framework of knowledge which makes it possible for them to interpret the many and various outside signals they are receiving and thereby makes them sensitive to relevant novel information.*

At least one way of achieving this is through the development of scenarios which are so designed as to enhance the decision-makers' understanding of the future. These must be rich in description, rather than contain numbers alone; their prime purpose is for use in debate. The scenarios must also be relevant to those issues perceived as the most important by the decision-makers.
All this can only be of use if the expert, the planner, accepts his role not as the soothsayer or prophet but as an adviser whose primary objective is to enhance and improve the manager's thinking. In no way does this reduce the value of modelling and of econometric models. They remain powerful tools of analysis and of research. Their relevance, however, must be judged not against the criterion of how far they mirror the unknowable reality of the future, but how far they enhance the knowledge and understanding of decision-makers.

COMPLETE OTA REPORTS

A Checklist of Components Sufficient to Guarantee Completeness

1. A list of congressional action options?
Test: by inspection.
Manageably small?
Test: by inspection.
Objectively obtained?
Test: the soundness of the assessment report's argument for the objectivity of the paring method used to reduce its size.
Jointly exhaustive?
Test: a) if a checklist is provided, are there any obvious omissions? b) if a stakeholder survey is used, were the stakeholders representatively sampled? c) is the assessment report's own argument for the exhaustiveness of the action options sound?
Feasible?
Test: Logically possible? physically possible? socially permissible?
Logically possible?
Test: a) if the congressional options are clear, are any contradictory? b) if the options are not all clear, are the inconsistent subsets identified?
Physically possible?
Test: a) are the options obviously physically possible? b) alternatively (and/or) to (a), was the list of options examined and judged plausible by appropriate experts? Is every known and significant scientific dispute about the plausibility of any option reported?
Socially permissible?
Test: legally permissible? Not morally unacceptable?
Legally permissible?
Test: consistent with legal precedents?
Not morally unacceptable?
Test: either by inspection or by means of public polling.
Objectively tested?
Test: a) was a representative sample of stakeholders polled? were stakeholders among the inarticulate sectors of society and among the traditionally unrepresented academic disciplines polled? b) does the assessment team report their criterion of acceptability? is it sound? c) alternatively (and/or) to (a) and (b) above, is it obvious that the options are not morally unacceptable?

2. A set of scenarios?
Test: by inspection.
Mutually exclusive?
Test: by inspection.
Relevant?
Test: either by inspection or by means of a survey of outcome's desirabilities. a) if a survey was used, were stakeholders representatively sampled? b) alternatively (and/or) to (a), are the features in the scenarios obviously relevant?
Manageably small?

Test: by inspection.

Objectively obtained?

Test: the soundness of the assessment report's argument for the objectivity of the selection method.

Practically exhaustive?

Test: a) did any technique used to construct the set of relevant features have a known and currently correctable defect which could result in overlooking a relevant feature? b) was any technique that seems likely to contribute to the set of relevant features overlooked? c) was a representative sample of stakeholders surveyed? d) are the assessment team's criteria for choosing features of the world to present for stakeholder consideration discussed? are they likely to ignore relevant features?

Were all the relevant physical, biological, economic, social and political effects identified by the assessment team?

Test: does the assessment report argue persuasively that expert consensus in each area is that no further relevant effects are known?

Were all relevant natural and institutional states that are not effects identified by the assessment team?

Test: does the report argue persuasively that no further relevant features are identifiable?

3. A set of unaggregated desirability polls, one for each outcome?

Test: by inspection.

Stakeholder opinions?

Test: a) if stakeholder polling was used, was a representative sample polled? are the results summarized according to stakeholder characteristics provided by Congress or known to be useful to Congress? are the results presented by congressional district? b) alternatively (and/or) to (a), are the desirabilities of the outcomes obvious to the appropriate legislators? If this is not clear, does the assessment team argue persuasively that it is true?

4. A set of numerical conditional probabilities, one for each outcome?

Test: by inspection.

Objectively obtained?

Test: a) are observed relative frequencies directly available? b) alternatively (and/ or) to (a), were the probabilities derived from observed relative frequencies supplemented by testable theories? c) alternatively (and/or) to (a) and (b), were the probabilities obtained from simulation models? d) alternatively (and/or) to (a), (b, and (c), were the probabilities obtained from expert testimony?[35]

[35] from: Lewis Gray, 1982, "On 'Complete' OTA Reports", *Technological Forecasting and Social Change* 22, 3 and 4, pp. 299-319.

CROSS IMPACT ASSESSES CORPORATE VENTURES, Anoniem, Chemical & Engineering News, 16 april 1973, pp. 8-9

Monsanto uses analytical technique to help management sort out alternative business strategies and policies

You have a promising new polymer fresh from the laboratory. Your sales manager is enthusiastic, he sees a potential 300 million pound-a-year market by 1980. The engineering staff estimates that a plant could be on stream in two years at a cost of $50 million. The purchasing department foresees no difficulty in obtaining raw materials at a reasonable price. A customer evaluating the product has worked out an interesting application for packaging. On the other hand, the impact on new plastics of impending waste disposal regulations is unsettled. Your financial vice president has doubts about the availability of capital to commercialize the product. You've heard that a competitor is developing an unrelated material that appears to be aimed at the same market. And your enthusiastic customer may be the target for a takeover bid from another resin maker.

How do you balance all this complexity of issues and conflicting judgments to decide the best approach to marketing the product - or even if it is a worthwhile commercial gamble at all?

For the past three and a half years, Monsanto has been tackling problems of this type with a technique called cross-impact analysis. Cross-impact analysis - actually a family of related techniques - was initially developed for technological forecasting by Olaf Helmer and Theodore J. Gordon at the Institute for the Future, a forecasting oriented think tank, from a simulation game they had designed in 1966 for Kaiser Aluminum. Its strength lies in its ability to assess, in terms of changing probabilities or timing, the interrelationships and interactions among a broad series of possible future events or policies, and to spotlight those that are most critical. Statistical analysis (with a computer) of the cross impacts throughout an array of these key events brings into focus those that seem likely to have the greatest influence on a planned endeavor and also gives insight as to its probable success or failure.

J. Kenneth Craver, Monsanto's inhouse futurologist, has taken the technique and adapted it for project evaluation and the sorting out of alternative corporate strategies regarding new products, plant investment, diversification and reorganization. Unlike most other cross-impact methods, which generally look at events occurring at some discrete slice of time in the future, Monsanto's version evaluates changes in probability or timing over an extended period of time. Now Monsanto has licensed the computer programs for its dynamic model to Futures Group, another future-oriented consulting firm which plans to offer it to its industrial clients.

"Cross-impact analysis is not a crystal ball," cautions Theodore Gordon, now president of Futures Group. "You can't just plug in numbers and expect to get a firm answer. But it can provide insights into the interactions of various future events on one another, even events like social or regulatory changes that can't be quantified."

"It's not as precise as conventional cost-benefit analysis," adds Monsanto's Ken Craver. "But because it lets you look at things that more quantified methods can't handle, it adds a degree of reliability that you don't get if you depend solely on cost-benefit analysis."

The method is a powerful device for "forcing decision makers to be a bit more thoughtful about what goes into their planning," according to Mr. Craver. "It makes them treat subjective or intuitive value judgments in a logical manner, consider options that they might not otherwise have thought of, and recognize the consequences of making or not making a decision. There is no question that it has prevented us from making mistakes in launching new products and led us into moves different from what we had originally planned".

Dr. Constantine E. Anagnostopoulos, head of Monsanto's New Enterprise division, has participated in several cross-impact exercises and is equally enthusiastic. He plans to run an analysis on every major project his division undertakes. And he thinks it makes sense to repeat the analysis several times during the commercial evolution of a new product - whenever major premises change or basic strategies are shifted. "The actual events you look at may not change," he notes, "but their importance or their relationships to one another may be different. And if you come up with answers that are unexpected or different from the previous time around, that may be a signal of trouble ahead."

At Monsanto, a cross-impact analysis starts by getting together a team of five to eight people concerned at the decision making level with the project under study. These might include, for example, the head of the division the project manager or coordinator, and top representatives from research, production, commercial development, and marketing. "You need the experts on the team," Dr. Anagnostopoulos stresses, "because you have to have concrete inputs rather than guesses."

Each member of the team is asked to list in advance the critical events, policies, or objectives that will have a bearing on the project's success or failure. As a practical matter, the analysis can handle only 25 to 30 such events at most. If a greater number are suggested therefore, the list is screened, refined and consolidated to a manageable total. "One of the chief benefits of the analysis," Dr. Anagnastopoulos asserts, "is that it makes everyone do a lot of homework in identifying the key issues in a well-defined manner."

The analysis zeroes in on the problem areas: Where and how big are the potential markets? How large should the initial plant be? What will competitors do in response? When will we get satisfactory return on our investment? How might future technological developments affect the venture? What type of government regulation or environmental pressures can we expect?

Once a list of crucial events has been put together, the team is assembled for a strenuous, full-day evaluation. First, it must forecast collectively for each given event the year when the probability of its occurring is 20% and when the probability is 80%. The events also are ranked in order of occurrence, since some may be impossible before others have taken place. For instance, the product cannot be commercialized before a plant is operating, and the plant cannot be built until money for it has been appropriated.

The group is next asked to assume that one of the events actually occurs (its probability now becomes 100%) and to estimate the relative positive or negative impact of that occurrence on every other event. A capital investment for a plant may greatly enhance the probability of reaching a sales target, for example; licensing technology abroad may have no effect on domestic sales, introduction by a competitor of a better or cheaper product may inhibit sales severely. The team, as individuals, votes rapidly on these impacts, using a scale of values ranging from highly positive to highly negative. As the team works through the entire list of events in this fashion a matrix of assessments (based on the team's median vote) is built up. In an analysis involving 30 separate events, a matrix of 870 cells must be filled in, an exercise requiring six to eight hours of concentrated effort. "It's an introspective, soul-searching activity," Mr. Craver points out. "At the end of a day of thrashing out all the interactions, tensions are high and everyone is angry with everyone else. But even before the results are evaluated, the pivotal events affecting the project probably will stand out clearly to everybody who has participated."

The day's results are fed into a computer, using a Monsanto-developed program that generates two growth or probability curves for each event. One of these is derived from the probabilities assigned initially to the timing of the event; the second indicates its probability as adjusted on the basis of judgments of the impact on it stemming from the occurrence of every other event.

If the two curves lie close together presumably the team's original estimates and strategies are consistent with the analysis and thus probably "correct," at least insofar as they relate to the events included in the overall exercise. If the second curve is shifted away from the first, the particular event should be scrutinized more carefully. The matrix itself may be internally inconsistent or lacking certain critical events. On the other hand, the participants may be too optimistic or too pessimistic in judging the probability (or desirability) of the event. Or perhaps they have misunderstood the nature of the event itself or are victims of self-delusion, misconception, or inadequate planning.

"Shifts in the probability curves suggest a change in strategy may be in order," says Ken Craver. "It may be that the team hasn't really thought its plans through or it doesn't really believe in the approach it is taking. The cross-impact analysis brings this out and lays it on the table early."

"You may not get exact, quantitative results from a cross-impact study," adds Dr. Anagnostopoulos, "so that people who like to deal with numbers will be disappointed in the output. But the technique does expose prejudices and differences of opinion. And it detects fuzzy thinking and inconsistencies. Perhaps you haven't given enough thought to what your competitors will do or considered all the alternative approaches. What you are doing in cross-impact analysis is putting yourself in the position of an outsider, so that you can stand apart from the project and wear the two hats of critic and advocate at the same time."

THE ANALYTIC HIERARCHY PROCESS (AHP), in: Forecasting and management of Technology, Alan L.Porter, A.T.Roper, T.W.Mason, F.Rossini, J.Banks, 1991, John Wiley & Sons, Inc. New York, pp. 363-366.

AHP was created by Thomas Saaty to structure complex judgments (Saaty, 1980; Saaty and Kearns, 1985). It does this through four basic stages:

1. Systematizing the judgments into a hierarchy or tree
2. Performing elemental, pairwise comparisons
3. Synthesizing those pairwise judgments to arrive at overall judgments
4. Checking that the judgments combined are reasonably consistent with each other

The AHP process is hierarchical. As an illustration, assume that your objective, the highest level of the decision tree, is to get a good job. Suppose you break this objective down into three criteria, constituting a second level of the tree. Assume these criteria are salary, location, and opportunity to advance. The importance of these three criteria relative to each other can be determined using the AHP procedure. Furthermore, suppose that five alternative jobs, each in a different city, are being considered. These can be compared for each of the three criteria using the AHP procedure at this level (yielding local priorities for the set of elements on the second level immediately above). AHP's hierarchy is structured from the top down, much like relevance trees.
Once the decision hierarchy is specified, you can turn to the judgments to be made. People can judge between two items more easily than they can make composite judgments of multi-ple items all at once. Therefore, AHP uses pair-wise comparison among each relevant pair of items as the basic judgments. Other techniques, such as interpretive structural modeling (ISM), also use pair-wise comparisons. However, in contrast to ISM's dichotomous judgment, AHP employs a nine-point scale. Saaty (1980) documents the superiority of this scale over alternatives.
Consider a sample judgment as to the relative importance of two of the good job criteria of the example just introduced. How much more important is salary than location? Referring to the Table, suppose you feel that salary is more important, meriting a ranking of 4. AHP will fill in the complementary judgment of location compared to salary with the reciprocal value - 1/4. Concerning the three criteria of salary, location, and opportunity, you will be required to make two more judgments (salary versus opportunity and location versus opportunity). AHP will fill in the complements. The result will be a 3-by-3 matrix (with 1's on the diagonal - that is, salary is equally important to salary). The same items appear as the rows and as the columns of the matrix - salary, location, and opportunity.
The next AHP stage is to synthesize the judgments within a given matrix (for local priorities) and then across matrices (global priorities). The idea is quite simple - to collapse the set of separate judgments into a properly weighted overall judgment.

Importance	definition	Explanation

1	Equal importance	Alternatives contribute identically to the objective
3	Weak dominance	Experience and judgment slightly favor one alternative over the other
5	Strong dominance	Experience and judgment strongly favor one alternative over the other
7	Demonstrated dominance	One alternative's dominance over the other is demonstrated in practice
9	Absolute dominance	Evidence favoring one alternative over the other is affirmed to the highest possible order

Relative importance scale
Source: Based on Saaty, 1980.

Calculation involves matrix mathematics but need not be a direct concern. The TOOLKIT, or more elaborate programs such as EXPERT CHOICE, provide these weighted priorities for each matrix.
The fourth AHP stage is to check the consistency of the judgments in each matrix. Collections of pairwise judgments are apt to show inconsistencies. These may reflect crude scaling (such as A seems a little better than B; B seems a little better than C; C seems a lot better than A - the imprecision of the "little better" designation leading to considerable uncertainty). Or raters may just be flagrantly inconsistent (for instance, preferring A to B, B to C, and C to A). AHP provides a helpful indicator to signal the degree of inconsistency in a matrix of judgments. This requires extension of the synthesis calculations.
Calculation of the degree of inconsistency again requires matrix manipulations. These yield the Consistency Ratio:

$$\text{Consistency Ratio} = \frac{\text{Consistency Index}}{\text{Random consistency number}}$$

The random consistency number indicates an expected value if judgments were taken at random over the scale from 1/9 to 9. The random consistency number varies as a function of the size of the element set:

matrix size	2	3	4	5	6	7	8	9	10
random consistency	0	0,58	0,90	1,12	1,24	1,32	1,41	1,45	1,49

Saaty suggests that the Consistency Ratio should be 10 percent or less: sometimes up to 20 percent may be tolerated. The box discusses practical steps in conducting an AHP analysis. These integrate the four stages into a typical sequence of activities.
Saaty and Kearns (1985) document a range of AHP applications that illustrate the following:

- Inclusion of interdependencies among criteria and how these alter priorities in comparison to assuming independence of the criteria
- Formulation and comparison of alternative scenarios, using an example of seven scenarios for higher education in the United States through 2000, analyzed over four primary factors (economic, technological, etc.), six actors (faculty, government, etc.), and various actor objectives (four faculty objectives, six governmental objectives, etc.)

Steps in the Analytical Hierarchy Process (Based on Saaty and Kearns, 1985).

1. Define the problem and what you want to know. Uncover assumptions and preconceptions reflected in the problem definition; revise the problem definition if these are not viable. Identify affected parties; check how they define the problem. Consider ways for them to participate in the AHP.

2. Structure the hierarchy from the top - that is, from the overall objective to the intermediate level(s) factors or criteria to the lowest level (usually the alternatives under consideration). Check that levels are internally consistent and complete and that relationships between levels are clear.

3. Construct one pairwise comparison matrix covering the set of elements in the lowest level for each element in the level immediately above. In complete simple hierarchies, every element in the lower level affects every element in the higher level. In other hierarchies, lower-level elements affect only certain upper-level elements, requiring construction of unique matrices.

4. Make the judgments to fill in the matrices $-n(n-1)/2$ judgments per each $n \times n$ matrix. The analyst (or the group participating) judges whether element A dominates element B - if so, inserting the suitable whole number (see Table) in the cell at row A, column B - or, if B dominates A, inserting the whole number in row B, column A. The reciprocal is automatically inserted in the counterpart cell.

5. Calculate the Consistency Ratio for each matrix. If unsatisfactory, redo the judgments. Repeat steps 3 through 5 for all levels of the hierarchy.

6. Analyze the matrices (preferably using a computer program such as the TOOLKIT or EXPERT CHOICE) to establish local and global priorities. Check the hierarchy's consistency by multiplying each Consistency Index by the priority of the corresponding criterion and adding them up; then compute a consistency ratio. If this is too high, redo the judgments (for instance, rephrase questions or recategorize elements). Saaty recommends that each set include no more than seven elements; larger sets can be broken down into multiple groups, repeating one element in each to use as an anchor.

TREND EXTRAPOLATION, in: FORECASTING AND MANAGEMENT OF TECHNOLOGY, Alan L.Porter, A.T.Roper, T.W.Mason, F.Rossini, J.Banks, 1991, John Wiley & Sons, Inc. New York, p. 169-175, 185-187.

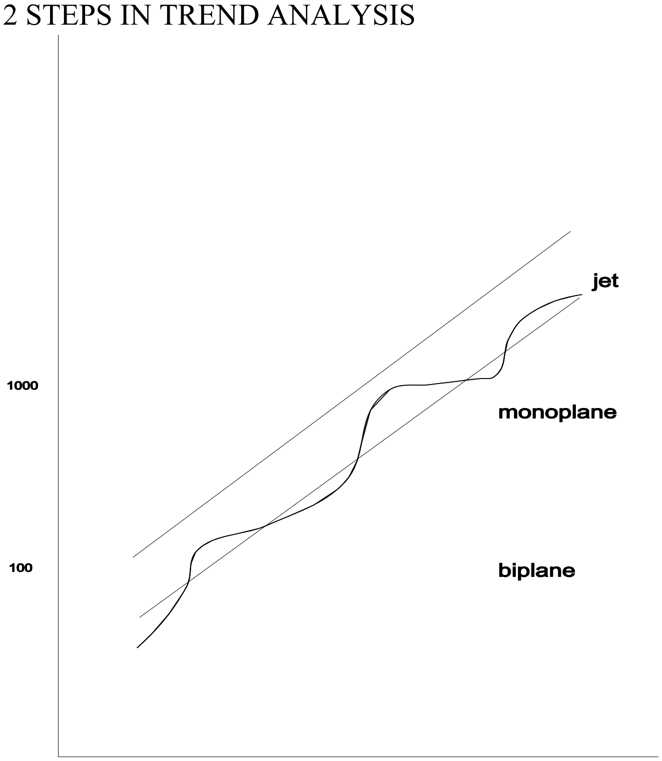
OBJECTIVES

This chapter explains and illustrates the use of the most applicable forms of trend extrapolation for technology forecasting and presents trend analysis as a four-step process. The chapter emphasizes two key growth models, Fisher-Pry and Gompertz. Finally, it illustrates how the Lotka-Volterra equations offer a promising general framework for trend modeling.

1 TREND ANALYSIS IN TECHNOLOGY FORECASTING

Technology forecasting relies largely on naive (direct) time series analysis. This implies major assumptions about the nature and permanency of both context and structure. Trend analysis methods can yield valid forecasts when supporting and competing mechanisms in the larger environment remain constant over the time horizon of the forecast or when changes in these mechanisms cancel one another. Even under those appropriate conditions, trend extrapolations should be used in conjunction with complementary technology forecasting methods, especially expert opinion and monitoring. This chapter stresses extrapolation over time. It applies regression techniques to fit selected nonlinear relationships that are especially suited for technology forecasting.

2 STEPS IN TREND ANALYSIS



Enveloppe curve.

Once the variable(s) have been chosen and the necessary data have been obtained, trend analysis can begin. The textbox outlines the basic steps. Step 1, box, model identification, draws upon insight into typical patterns of technological innovation, solid knowledge about what driving the change and empirical evidence. S-shaped growth should be considered the most likely form. Section 3 contrasts the Fisher-Pry and Gompertz models - two approaches that produce S-shaped curves. However, other models sometimes merit serious consideration as well.

Exponential growth often holds over certain periods, or epochs (Hamblin, Jacobsen, and Miller, et al.. 1973). The growth rate then shifts and another epoch emerges. Over multiple epochs, continuing exponential growth becomes apparent, but succeeds at different rates. Exponential growth may hold for the time period of interest, but possible physical or social limits that could slow or stop growth must be anticipated.

Steps in Trend Analysis

1. Identify the proper model. Prominent alternatives include:
    a. S-shaped growth curves
        1. Fisher-Pry (or equivalently, Pearl)
        2. Gompertz
        3. Technical progress function
    b. Learning curve
    c. Exponential growth
    d. Linear
2. Fit the model to the data
    a. Graphically
    b. Solve for the constants in the equation
3. Use the model to project
    a. Graphically
    b. Mathematically
4. Perform sensitivity analysis and interpret the projections.
    a. Compute confidence intervals
    b. Consider outside factors.

An envelope curve can be constructed by stacking S-shaped curves, one after and over another. The figure shows a classic example of how a series of technological developments, each S-shaped, can combine to drive a parameter forward, possibly along an exponential frontier. As one technology approaches a peak, R&D provides a successor that fulfills the appetite of a still-hungry market. An envelope curve could be plotted through the peaks and the valleys of the successive technologies to depict the general trajectory of the development pattern.

The technical progress function is another important model. It measures growth as a function of effort instead of time. The notion is that progress in developing a technology starts slowly as many impediments must initially be overcome, advances rapidly for a period, and then slows as the easy improvements are "mined out." This is, of course, the S-shaped growth curve in another guise. The tapering off of technical progress for constant increments of additional effort implies dwindling productivity. The R&D organization receives the greatest payoff from its effort during the steep portion of the technical progress curve. After that, marginal returns per unit of effort diminish. The technical progress function provides a vital signal to those who would manage R&D for a technology - that is, at some point continued investment in R&D will deliver less and less. This means that the natural momentum of continuing to do what has delivered good results in the past, must be stopped - that large, successful research group must be reassigned to fresh tasks or R&D productivity will drop.

A related notion is the learning curve, which addresses the improvement in productivity often seen as a technology production process matures. This is conventionally depicted as a power function:

$$Y = aX^{-b}$$

or equivalently,

$$\log Y = \log a - b \log X \qquad (1)$$

where Y is the number of direct labor hours required to produce the Xth unit: a is the number of direct labor hours required to produce the first unit: X is the cumulative number of units produced (not time); and b is a parameter that measures the rate that labor hours are reduced as cumulative output increases (Argote and Epple, 1990).

Analysis of the learning curve for a technology can help predict declining production costs. The decline is attributable to organizational learning as personnel gain experience in refining the production process - in other words, the learning curve is due to both technology and people factors. Each increase in cumulative output leads to a reduction in unit production cost. While this "progress ratio" varies greatly from technology to technology, the modal value is about 80 percent - that is, each doubling of cumulative output leads to a 20 percent reduction in unit production cost (Argote and Epple, 1990).

Projection of a learning curve can help a technology manager monitor his or her own production processes to ensure that learning is progressing reasonably. Projection of a competitor's learning curve could help gauge whether to enter a market.

Having identified a promising model, the second step is to fit the model to the data. Should that fit prove poor, the forecaster may wish to reconsider the choice of model. The choice of a model should never be made by fishing through a grab bag of models and picking the one that fits the data best. Noisy data can mask true relationships; there is no substitute for solid conceptual underpinning.

Begin by graphing the data. Remember that individual data points may be problematic. This is important as outlier points can exert great influence over mathematical curve fits. Outliers may result from special circumstances, mistaken measurement, transcription errors, or just a divergent value that cannot be ignored. The forecaster may want to examine various transformations. For many forecasting purposes, fitting a line to properly transformed data will provide a satisfactory basis for extrapolation. Even if the equation for the trend is eventually calculated, the graph will provide an excellent check. It is easier to detect a bad extrapolation on a graph than from an equation.

Once the equation is determined, an extrapolation can be made graphically and/ or mathematically (Step 3). Next, perform a sensitivity analysis for the extrapolation (Step 4).

Calculation of confidence intervals provides vital information on the range of future values to be expected. Consider also what outside influences are important - that is, which ones could alter the trend substantially and that are reasonably likely to occur. This sensitivity analysis may be quantitative and/or qualitative. Cross-impact methods can be of use, and expert opinion can be quite helpful-at this juncture. A strategy that often proves highly effective includes these steps:

- ¨ Show your trend extrapolation to selected experts for their reactions.
- ¨ Raise specific questions about the external influences identified to see if the experts agree that they are likely and how they would alter the trend.
- ¨ Ask the experts to identify other factors likely to alter the trend.

After the sensitivity analysis has been performed, the projections should be interpreted. Forecasters often consider their job to be done when they provide trend(s), but this is not so. The implicit knowledge that has been gained in determining that trend should be made explicit. The forecaster needs to indicate

- ¨ Why a certain model has been selected
- ¨ How strong or weak the data are
- ¨ What factors are likely to interact with the trend and how probable they are
- ¨ How much confidence he or she has in the trend

Wherever possible, an open dialogue with the intended users will add significant value to the trend analysis.
(...)

3.3 Choosing Fisher-Pry or Gompertz

The model chosen by the forecaster must embody underlying characteristics that reflect those of the process to be forecast. This is not always easy: the process often is not well understood. However, the Fisher-Pry and Gompertz models offer an opportunity to illustrate model selection. The following discussion reflects the general concerns of model selection and of differention between growth and mortality models.
Equation 1 shows that Fisher-Pry, a growth model, assumes that the rate of change of the process is proportional to both the fraction of the market penetrated by the technology and the fraction that remains to be penetrated. It depends on the number of uses for which the new technology has been applied and on the number for which it is yet to be applied. This is analogous to the process followed by the diffusion of a new, technically advanced product. In such cases, knowing someone who owns one helps prospective buyers assess the technology's potential for them. Since the availability of the technology and of spares, repair facilities, and advice normally grows with the number of units in the field, diffusion is further enhanced by the sale of additional units. Thus penetration of the market is given, not only by potential sales, but also by the sales that have been achieved.
In the diffusion process modeled by Fisher-Pry, initial sales of a new technology are difficult despite its promise and the size of the potential market. This is true largely because the technology is unknown and unproven. Moreover, in the early stages, field support is likely to be poor, there may be institutional barriers to overcome (such as, licensing), and survival of the technology and its supplier is uncertain. Thus adoption implies considerable risk. As applications grow, so does general knowledge of the technology and its advantages. The support infrastructure improves, as does confidence that the purchaser will not suffer either from inadequate support or from the financial failure of the manufacturer. From this point, penetration grows rapidly. At some stage (for example, after 50 percent penetration), further

penetration becomes increasingly more difficult because it often involves sales to companies that may not benefit as greatly from the technology or that have marginal capability to finance adoption. Thus the rate of penetration slows.

The Gompertz model embodies quite different dynamics. For example, although not readily apparent from Equation 5, for penetrations greater than 50 percent, the rate of penetration depends primarily on the fraction of the market remaining. Thus the Gompertz model is appropriate for forecasting market penetration by technologies for which initial sales do not make subsequent sales easier. This dynamic usually is found when a new technology offers no clear-cut advantages over an old. In such instances, an older technology is replaced by a newer technology that performs the same tasks with essentially the same financial and/or functional efficiency. Purchases simply replace equipment that has worn out or has been destroyed. It is this characteristic of the Gompertz model that leads to its classification as a mortality model.

Since the dynamics of the two models are different, it is reasonable to expect significant differences in the forecasts they produce. For example, the Fisher-Pry model forecasts a more rapid penetration than does the Gompertz. However, either model can be made to fit a handful of initial data points. Thus goodness of fit cannot be used as a criterion to determine which model to use; instead, the dynamics of the model must be matched to the dynamics of the process being forecast.

Not all substitution processes are clear-cut. Some, for example, appear to be driven by a mix of new technology diffusion and old technology mortality. For example, some owners of prestige automobiles may replace them with newer models that offer significant technological improvements (such as anti-lock braking systems or fuel injection). Others may replace to 'keep up with the Joneses' even when no clear technological superiority is offered. Still others may be driven (no pun intended) to replace worn-out units by nearly identical ones because of the prestige and tradition of an established design ("only the very best people own a ...").

In cases in which the underlying dynamics of the penetration are unclear, the best guide is to consult those whose job it is to understand the dynamics of the market. The overly optimistic forecasts of demand for innovative technology products, such as home computers and expert systems, for example, can be partly attributed to the failure of forecasters to consult such experts (Wheeler and Shelley, 1987). However, Schnaars (1989) notes that such experts often get caught up in the "zeitgeist" (prevailing wisdom) of the times and tend to be overly optimistic. Historical analogy to similar situations also can be helpful.

The forecaster must match the dynamics of the model to those of the process. To reiterate, the dynamics of the Fisher-Pry model are appropriate to cases of technology diffusion: the dynamics of the Gompertz model are appropriate to cases of replacement driven by equipment deterioration rather than technological advantages. Occasionally, as Lenz (1985) notes, industries have been able to match equipment deterioration to new technology innovation; however, the difference in the rates associated with each mitigate against achieving that balance often. When the dynamics of the process are unclear, consult the experts, compare alterative models, and pray.

3.4 Selecting an Upper Bound for the Forecast

The preceding discussions, have emphasized forecasting the market fraction, $f = Y/L$, where L is the upper bound for Y. This approach simplifies the presentation; however, it also disguises the fact that L is a third parameter that must be estimated to employ forecasts made with the Fisher-Pry and Gompertz models. An accurate estimate of L is important for a number of reasons. First, L must be known to formulate the time series data for forecasting. Second, an incorrect upper bound can seriously distort values of the fitting coefficients (b and c in the Fisher-Pry model) and hence the subsequent forecast. Further, the technology manager generally needs a forecast, not only of the fraction of the market that will be penetrated, but also of the number of units that will be sold.

In the case of Cable TV, the upper limit was set as the number of households with TV sets - a number that grows with time and therefore, is itself a subject for forecast. The range of *f* for Cable TV could be established as zero to one with some confidence. In other instances, this cannot be done so easily (for example, in forecasting the upper limit for the sales of subcompact automobiles, the total number of automobiles sold would not be an accurate upper bound). Nor is the limit for the functional capacity of a technology (for example, the precision of a manufacturing method or the level of concentration of a chemical compound that can be detected) so easily set.

In some instances, forecasters have used the data to establish the upper limit (for example, in the Fisher-Pry method, data fits would be performed for b, c, and L). As Martino (1983) notes, this is bad practice because initial data are relatively insensitive to the upper bound and thus do not provide reliable guidance. Rather, L should be set by natural or fundamental limits to the process. This, of course, requires knowledge of the technology and the market place that the forecaster may not possess. Therefore, in cases where the upper bound is unclear, the forecaster should work closely with experts in the field to determine a reliable estimate.

CROSS-IMPACT ANALYSIS, in: FORECASTING AND MANAGEMENT OF TECHNOLOGY, Alan L.Porter, A.T.Roper, T.W.Mason, F.Rossini, J.Banks, 1991, John Wiley & Sons, Inc. New York, p. 223-228.

A basic limitation of many forecasting techniques is that they project events and/or trends independently (Stover and Gordon. 1978); thus they fail to account for the impact of events or trends on each other. For example, a successful nuclear fusion process could have a major effect on petroleum exploration. Likewise, the scarcity of petroleum resources holds great economic implications for the development of nuclear fusion. These two technologies do not exist in isolation. Each has a history; each is affected by developments in the other.
One approach to capturing interactions between events is to construct a model, that is, a formal representation of interactions among significant variables. There are several types that can be employed. A mathematical model uses equations to represent the system in which the events occur. Such models often require a major time investment to construct. Even with this investment, model coverage usually is limited (for example, mathematical models of inventory systems, of the economy, or of resource allocation systems). There are, however, special models that cut across disciplines and account for the effect of one event upon another. In the technology forecasting arena, one such model is cross-impact analysis (CI). Basic CI concepts are widely used and have applications in many areas, including natural resource depletion, strategy and tactics for warfare, institutional change, organizational goals, communication capability and computer capabilities.
Since CI deals with the future, it involves uncertainty. Therefore, it is a stochastic rather than a deterministic model. Traditional CI is focused on the effects that interactions between events have on their probabilities of occurrence. Thus it deals with discrete events and incorporates no dynamic (time) dimension. While still discrete, the dynamic dimension can be added to CI using the concepts of Markov chains. This modification of traditional CI can be employed to study the chain reaction of events/trends on other events/trends over time. These CI approaches are described in the following subsections.

2.1 Traditional Cross-Impact

The concept of CI arose from a game called 'Future' which Gordon and Helmer devised for Kaiser Aluminum in 1966 (Helmer, 1983). In the game, a future world was constructed in which some or all of 60 events might have taken place (technological breakthroughs, passage of legislative measures, natural occurrences, international treaties, etc.). Each event was assigned an initial probability of occurring. As play progressed; these probabilities changed. Part of the change was due to actions of the players, the remainder was determined by the occurrence or nonoccurrence of other events. Change of the latter type gave rise to the concept of CI.
A specific example is useful to understand how traditional CI works. Suppose we are planning for a particular communication technology, say facsimile transmission (fax). Fax technology allows text and images to be transmitted over normal telephone lines. Hard copy input is provided by the sender: it is then transmitted, providing hard copy at the receivers end. Transmission is fast. 15 seconds or so per page: thus it is much cheaper than conveying the same information verbally (even when that is possible) and much faster than using the mail service. We wish to know what the future of fax technology will be over the next five years. First, we must determine the ways in which the operating environment is likely to change over that time horizon. Then we must identify events that could have noticeable

impacts on the use and/or on the planned uses of fax and the probabilities that these will occur.
Suppose the events are identified as E1, E2, E3, ... Em. These represent entirely external determinants - that is, natural or man-made events over which we have no control (such as a global economic depression or legislation imposing a large tax on each fax received). Events completely under our control are not included. These must be treated differently. If the number of events grows too large, it may be necessary to rank them and retain only those that are most important. This could be accomplished by Delphi polling or by having interested parties assign points to each event on a scale of 0 to 100. For the example of fax transmission, suppose we have identified the four events shown in Table 1. For convenience, we have arranged our work space in an occurrence matrix with the events E1 through Em ordered both across the top and down the lefthand side of the array. The next step is to estimate the probability that each will occur. These estimates are called the marginal probabilities. (They also are sometimes referred to as ceteris paribus - all else equal - probabilities to indicate that they are estimated without considering any of the other events.) These probabilities are subjective and might be estimated through some Delphi-like procedure. For our example, Table 1 shows that we have estimated the probability that there will be increased taxes or costs on fax transmissions to be 0,70.

Table 1 Occurrence matrix for fax example

The probablity of this event becomes:	If this event occurs:			
	Increased taxes/costs	Negative legislation	Replacement technology	market saturation
Increased taxes or costs on transmission (0,70)*	1,00 P(1|1)	0,40 P(1|2)	0,70 P(1|3)	0,51 P(1|4)
Negative Legislation (0,40)*	0,20 P(2|1)	1,00 P(2|2)	0,38 P(2|3)	0,31 P(2|4)
Development of replacement technology (0,60)*	0,90 P(3|1)	0,72 P(3|2)	1,00 P(3|3)	0,33 P(3|4)
Market Saturation (0,45)*	0,33 P(4|1)	0,35 P(4|2)	0,05 P(4|3)	1,00 P(4|4)

* Initial marginal (ceteris paribus) probability.

We have completed two components of the CI matrix: the events critical to the forecast have been identified and their initial (marginal) probabilities of occurrence have been estimated. The cells of the matrix will be used to record the conditional probabilities (that is, the probability that event i occurs given that event j occurs). These probabilities are the heart of CI: they portray the impact that the occurrence of any event has on the probability that any other event will occur.
The conditional probabilities must be estimated next. First note, however, that the matrix diagonal entries all will be 1,00, for it is certain that event i will occur given that it has occurred. The first step is to compute the statistically acceptable range of conditional probability for each cell (pair of interactions) above the diagonal. These ranges will provide guidelines if we have no other basis from which to estimate the conditionals. This can be done using the marginal probabilities established previously for each event. To explain how

to compute this statistical range, we must first introduce some statistical notation.

P (i) = probability that event i will occur (the marginal probability of i)

P (i|j) = probability that event i will occur given that event j has occurred (the conditional probability of i given j)

P (_) = probability that event i does not occur

P (i/j) = conditional probability that event i will occur given that event j does not occur

P (i∩j) = probability that both events i and j will occur (the intersection of events i and j)

P (iUj) = probability that event i or j or both will occur (the union of events i and j)

By using the laws of conditional probability and the probability of compound events, Sage (1977) showed that limits exist to the range of statistically acceptable conditional probabilities. If the occurrence of event j enhances (increases) the probability that i will occur, then

$P(i) \leq p(i|j) \leq [P(i)/P(j)]$ (1)

On the other hand, if the occurrence of j inhibits (decreases) the probability that i will occur, then

$1 + \{[P(i) - 1]/P(j)\} \leq P(i|j) \leq P(i)$ (2)

Note that only the initial marginal probabilities P(i) and P(j) are necessary to compute these ranges, and they already have been estimated.
Next we must estimate a conditional probability for each of the cells above the diagonal and compare the estimates to the ranges computed from Equation 1 or 2. Estimates that violate the computed ranges should be retained if a solid rationale for them can be given. For example, in Table 1, the conditional probability P(1|2) has been estimated as 0.40. which is within the statistically acceptable range, 0.25 to 0.70. computed from Equation 2. However, if we had estimated that it should be 0.15 and had evidence to support our estimate, we would enter 0.15 instead. Alternately, we could elect to assign one of the extreme values of the range to such a probability. Thus, lacking strong evidence to support our estimate of 0.15, we might choose P(1|2) to be 0.25 instead.
Now that conditional probabilities above the diagonal have been estimated (the P(i|j)s), we can turn to those below the diagonal (the P(j|i)s). Here, we can use Bayes' rule to help. If the P(ilj) was in the range established by Equation 1 or 2, Bayes' rule says that the corresponding probability below the diagonal should be

$P(j|i) = [P(i|j)/P(i)]P(j)$ (3)

If P(ilj) was not in the ran,ge or if we do not agree with the value produced by equation 3, we should subjectively estimate the value of P(j|i). ln other words, if the values computed using

Bayes' rule are reasonable, keep them. Otherwise, estimate values believed to be more appropriate. For example, in Table 1, the conditional probability P(3|4) was estimated as 0.33, within the range of 0.11 to 0.60 computed from Equation 2. Therefore, Bayes's rule can be applied to give a value of P(4|3) = [P(3|4)/P(3)]P(4) = 0.25. Table 1 indicates, however, that we apparently had a strong rationale to support a lower estimate, 0.05.
Just as the occurrence of an event can affect the probability that another will occur, its non-occurrence can have an impact as well. In our fax example, for instance, if increased taxes or costs of transmission fail to materialize, then the impetus for and probability of replacement technologies will decrease. Thus we need to construct a nonoccurrence matrix (see Table 2).
Our last step is to estimate the entries for the nonoccurrence matrix, using the same philosophy as we did for the occurrence matrix. First we will compute the entries statistically from the following equation:

$$P(i|j) = [P(i) - P(j)P(i|j)]/[1 - P(j)] \qquad (4)$$

Lacking evidence to the contrary, these values will be entered. However, if evidence supports a different estimate, that estimate will be entered instead. Returning to the example

$$P(2|1) = [P(2) - P(1)P(2|1)]/[1 - P(1)] = 0.87$$

If we have no reason to estimate some other probability, then 0.87 should be entered into the nonoccurrence matrix.

Table 2 Nonoccurrence matrix for fax example

The probablity of this event becomes:	If this event does not occur:			
	Increased taxes/costs	Negative legislation	Replacement technology	market saturation
Increased taxes or costs on transmission (0,30)*	0,00	0,85	0,60	0,90
Negative Legislation (0,60)*	0,87	0,00	0,35	0,40
Development of replacement technology (0,50)*	0,48	0,52	0,00	0,78
Market Saturation (0,55)*	0,73	0,56	0,75	0,00

Initial marginal (ceteris paribus) probability of nonoccurrence P(i)=1-P(i).

Note that the diagonal entries in the nonoccurrence matrix will all be 0.00 since the probability of an event given that it has not occurred is 0. Negative probabilities predicted by Equation 4 should be set at 0, while predicted probabilities greater than 1 (certainty) should be set to 1.
The next stage in CI analysis is to simulate the effects of these conditional relationships. We must determine whether the initial estimates of event marginal probabilities are mutually consistent given these perceptions of how events impact each other.

If all the entries in the two matrices agree with results computed from Equations 1 through 4, then the initial marginal and conditional probabilities are mutually consistent. However, if one or more of the conditional probabilities differ from computed results, we will have to "play" the CI matrices to determine a consistent set of marginal probabilities. A computer-based Monte Carlo simulation can be used to do this:

1. An event is selected randomly (say Event 2 in Table 1).

2. A random number between 0 and I is generated and compared to the marginal probability of the event to determine if it occurs. Suppose the random number is 0.26, since $0.26 \leq 0.40$, Event 2 is assumed to occur. If the random number were greater than 0.40, it would be assumed that Event 2 did not occur.

3. The marginal probability of each remaining event is replaced by its conditional probability given that the event in Step 2 occurs or does not occur. That is, in our example P(i) is replaced by P(i|2) if Event 2 occurs, or by P(i|2) if it does not (i<>2). Thus, since Event 2 occurred in Step 2, the replacement values will be P(1) = 0.40, P(3) = 0.72, P(4) = 0.35.

4. A second event is selected randomly from those remaining (Events 1, 3, and 4), and Steps 1 through 3 are repeated. ln this play, the probability used in Step 2 is the value produced in Step 3 of the previous play. Thus, if Event 2 occurred in the first play and Event 4 is selected in the second, the probability of Event 4 used in Step 2 of the second play is P(4|2) = 0.35.

5. The process described in Steps 1 through 4 is repeated until all four events have been selected. All marginal probabilities are then returned to their initial values and the game is "replayed," typically 1.000 or more times.

6. Each time the game is "played" the events that occur are noted. The total number of occurrences divided by the number of games is taken as the final (marginal) probability for each event. The initial marginal probabilities are then replaced by the final marginal probabilities, which account for event interaction.

KSIM, in: FORECASTING AND MANAGEMENT OF TECHNOLOGY, Alan L.Porter, A.T.Roper, T.W.Mason, F.Rossini, J.Banks, 1991, John Wiley & Sons, Inc. New York, p. 241-246.

KSIM is a deterministic simulation model developed by Julius Kane (1972). KSIM extends the concepts of CI to produce a dynamic simulation that is easy to use yet sufficiently powerful to provide meaningful analysis of many real-world problems. The model retains the concept of the impacts of events on each other characteristic of CI. However, this concept is married to a differential equation that portrays an S-shaped (logistic) growth or decline of the variables being modeled. This equation provides the continuous, dynamical time-dependent) characteristics of KSIM. The logistic variation is a loose analogy to biological system growth. Since impact magnitudes are estimated subjectively, KSIM in effect utilizes both "hard" (objective) and "soft" (subjective) input. Thus it is an appropriate implementation of Kane's premise that experience, opinions, and judgments control decision-making.
The variables modeled by KSIM, X_i, are first identified, defined, and quantified. The maximum value of each variable is determined so that each can be normalized on a scale of 0 to 1. The initial value of each is also estimated. The simulation marches forward from these initial values a step at a time using the differential equation

$$\frac{dX_i}{dt} = \sum_{j=1}^{N} (\alpha_{ij} + {}_{ij}\frac{dX_j}{dt}) X_i \ln X_i \qquad \mathbf{1}$$

where

X_i = the variable described
N = the total number of variables considered
X_j = the impacting variables
α_{ij} = the long-term impact of X_j on X_i
$ß_{ij}$ = the short-term impact of X_j on X_i

The solution to this logistic CI equation is

$$X_i(t + \delta t) = X_i(t)^{P_i(t)} \qquad \mathbf{2}$$

where

$X_i(t + \delta t)$	= value of variable at end of the time period
$X_i(t)$	= value of X_i at the start of time period
δt	= the time period

and

$$P_i(t) = \frac{1 + \delta t(\sum inhibitingimpactson\ X_i)}{1 + \delta t(\sum enhancingimpactson\ X_i)} \qquad \mathbf{3}$$

$$P_i(t) = \frac{1 + 0.5t \sum_{j=1}^{N} [|I_{ij}(t)| - I_{ij}(t)]\ X_j(t)}{1 + 0.5t \sum_{j=1}^{N} [|I_{ij}(t)| + I_{ij}(t)]\ X_j(t)} \qquad \mathbf{4}$$

and

$$I_{ij} = \alpha_{ij} + \frac{_{ij}}{X_j(t)} [\, dX_j(t)/dt] \qquad \mathbf{5}$$

While the equations appear formidable, operationally the concept is relatively straightforward. One must estimate the impacts of the level of each event (that is, level = value of X_j) on all other variables. This is the α_{ij}, which is determined in much the same manner as impacts in Cl. Then, the impacts of the rates of change of each event (dX_j/dt) and the slope of the trend in X_j on the other events ($ß_{ij}$) are estimated in the same fashion. Once these impact magnitudes have been determined, a relatively simple computer program (such as that included in the TOOLKIT) can be used to solve the equations and perform the forecast.
The characteristics of KSIM are pretty much what would be expected of a logistic curve. For example, when the sum of the inhibiting impacts is greater than that of the enhancing impacts, the power $P_i(t)$ in Equation 2 will be larger than one. And, since $0 \leq X_i(t) \leq 1.0$, $X_i(t + \delta t)$ will be smaller than $X_i(t)$ Further, all else being equal, the larger the variable causing the impact, the greater the magnitude of that impact will be. Note also that a given value of $P_i(t)$ will have less effect on the magnitude of X_i if X_i is near either 0 or 1. This produces the S-shaped variation we expect of growth or logistic curves.
KSlM is one of the few dynamic models that can be constructed and used with relatively limited time and resources. The general procedure that a group of technology forecasters or managers would use is as follows:

1. Discuss the problem and agree on the scope and boundaries of the simulation such as level of aggregation, spatial boundaries, and time frame)
2. Identify, define, and label the important variables and determine their initial values, ranges, and maximums. Normalize each variable on a 0 to 1 range.
3. Structure the long-term and short-term impact magnitudes and array them in matrix form as in CI. Impacts that increase the size of a variable (enhance it) are positive, those that inhibit it are negative. Numerical values for magnitudes are proportional to the size of the impact. For example, if X_1 is not impacted by X_2, then α_{12}, and/or $ß_{12}$, will be 0. This work sometimes is cut in half by considering only short-term ($ß_{ij}$) or long-term (α_{ij}) impacts.

Run the model and refine the impacts, variable definitions and/or values until the outcome is satisfactory. Usually a base case is run and the output is compared to a similar situation or to theoretical behavior. The process is repeated as often as necessary to produce acceptable results.

The model can now run to investigate the effects of changing initial values or basic assumptions or of introducing new assumptions. In this way, alternative futures can be examined and forecasts and trade-offs can be determined. KSIM also provides for a very useful extension, allowing external events or policy decisions to be added to the model as variables in the CI matrix. This is done by formulating the impacts of, say policy options, as additional columns but not rows in the CI matrices. Thus a decision to invoke a policy option impacts the variables, although the option is not itself impacted by the variables. Using this approach, the decision maker can systematically investigate the effects of policy decisions on the behavior of the system.

KSIM should be viewed as a process as well as a product. The benefit of KSIM accrues from building the model, as much as from operating it and analyzing the results. Building the model provides the format in which a team can structure the discussion of a complex issue. In that format, experience, opinion, and judgment can be incorporated along with hard data. Further, a completed model allows alternatives to be quickly formulated and their consequences to be assessed. Thus KSIM can provide an environment within which the manager can study and learn about complex situations.

The process makes a number of assumptions that imply limitations as well. KSIM assumes that a satisfactory model can be devised and that the variables and their interactions can be accurately defined. lt also assumes that realistic bounds can be placed on the variables; that a growth curve adequately represents the change patterns being studied; and that opinions, experience, and other subjective information can be formulated mathematically. Equally important, KSIM assumes that the pairwise relationships portrayed by the matrices adequately represent true causal interaction, a much more complex situation. Finally, KSIM models a deterministic world; however, the technology manager, the forecaster, and the rest of us live in a probabilistic one.

KSIM Applied to Model Fax Transmission

To clarify concepts, consider the example of fax transmission. We will examine four variables:

1. Number of fax machines (irrespective of sophistication), N
2. Median cost of fax machines purchased, C
3. Number of pages of fax transmission, T
4. Cost/transmission (regardless of length), S

We also assume a single policy option - taxing fax transmissions. Suppose that impact magnitudes were estimated on a scale from 0 (no impact) to +3 (major impact) and that we have defined a major impact for each variable as one that causes a 10 percent change in the level of the variable. The short-term and long-term impact matrices that are estimated appear in the table. Note that these impacts are merely presented as representative for the purposes of this example; they are not careful estimates.

Note that the policy impacts are represented as long-term impacts only and that they are incorporated by adding a column to the long-term impact matrix. In the base case (no policy intervention) the model predicts that the number (N) of fax units in operation quickly approaches the maximum value. Other variables grow rather quickly as well. However, the median cost (C) of a fax unit grows slowly and then declines. Before this model is used for forecasting and decision making, it would be necessary to verify, insofar as possible, and modify the variable initial levels and impacts. This process might begin by setting year 1 as some time in the past and checking to see if model predictions track historical variable behavior. It is easy to see that even this simple check might be difficult because of problems associated with gathering the necessary historical data. These problems might cause us to

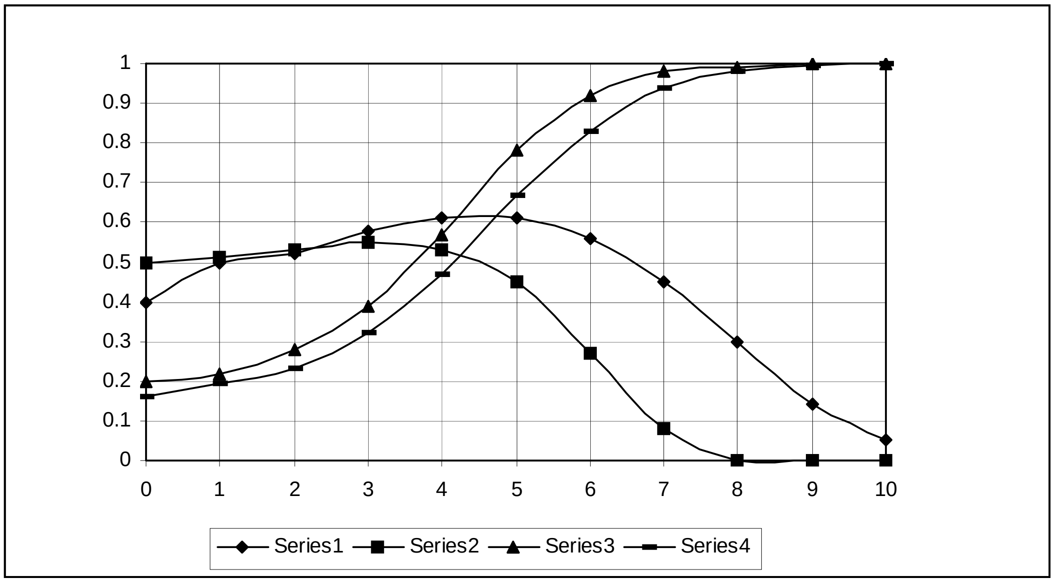
redefine variables to more readily fit the available data.
Assume the model has been fine-tuned and verified to our satisfaction and now we are interested in finding the changes in variable behavior that might be caused by a policy intervention. In the example, that intervention is a tax imposed on the normal cost of fax transmission. The impacts portrayed in the table for this policy option were constructed assuming that a tax of approximately 5 percent is imposed. We can model various degrees of policy intervention by choosing different values on the range 0 to 1 for the initial level of the policy variable. Note from the matrices that policies are represented by columns but not rows in the matrix. Thus policies impact, but are not impacted by, other variables. For this reason, the value of the policy variable does not change with time. We might choose a policy implementation of 0.2, for instance, to indicate that we will tax only those transmissions that involve documents of 10 or more pages.

TABLE 1 Long- and Short-term KSIM Impacts on FAX Transmission Variables

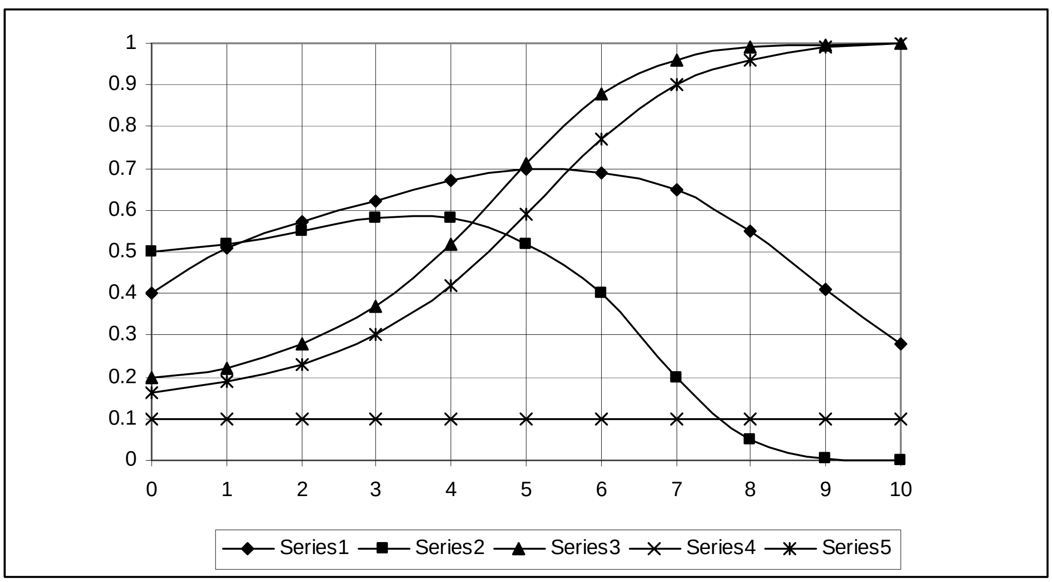
	Long-term impact of this variable					Short-term impact of the rate of change of this variable			
On this variable	N	C	T	$	Policy	N	C	T	$
N(0,2)*	3	-1	3	-1	-1	2	1	2	-1
C(0,35)*	-2	0	-1	0	0	0	0	0	0
T(0,15)*	2	0	3	-2	-1	2	1	2	-2
$(0,30)*	0	1	-1	0	1	0	0	-1	0

* Initial values: variables scaled from O to 1.

The policy intervention significantly changes the picture from the base case. Neither the number of fax units (N) nor the pages of transmission (T) rise as fast as before. The median unit cost (C) rises to a higher value before falling, but the peak is delayed about two years. Even with the tax, the cost per transmission ($) changes little over the first four years. However, it peaks at a much higher value about a year later than without policy intervention.
It is important to note that the policy impacts must be verified and modified just like the model itself. Only when the behavior seems reasonable to the manager or forecaster can a degree of confidence be placed in the results. This point is easily overlooked, for the model may assume a spurious credibility because of the sophistication of the computer equipment used to produce it. Beware!
The procedure could be extended to several policy options, or the policy option procedure could be used to model external events.



KSIM model of fax growth, base case.
(1=costs per transmission, 2= median cost machine, 3=number of machines, 4=number of pages transmitted)



KSIM model of fax growth, policy intervention.
(1=costs per transmission, 2= median cost machine, 3=number of machines, 4=policy, 5=number of pages transmitted)

Websites

United Nations University Millennium project	http://www.geocities.com/~acunu/
International Symposium on Forcasting	http://www.isye.gatech.edu/isf2001/
Plausible Futures	http://www.plausiblefutures.com/
Hudson Institute	http://www.hudson.org/
Sekretariat Zukunfsforschung	http://www.sfz.de/home/fs_home.htm
RAND Institute	http://www.rand.org/
World Future Society	http://www.wfs.org/
Institute for the Future	http://www.iftf.org/
Coates & Jarrett Inc.	http://www.coatesandjarratt.com/
Institute of Prosp. Techn. Studies, Sevilla (EU)	http://www.jrc.es/welcome.html
Int. Association of TA and Forcasting Institutions	http://www.IATAFI.ORG/
International Association of Impact Assessment	http://www.iaia.org/

Netherlands:

Stichting Toekomstbeeld der Techniek	http://www.stt.nl/
Advieraad Wetenschaps en Technologiebeleid	http://www.awt.nl
Wetenschappelijke Raad voor het Regeringsbeleid	http://www.wrr.nl/