1. General introduction: system innovation and transitions to sustainability

Frank W. Geels, Boelie Elzen, Ken Green

Modern societies face structural problems in several sectors. In the energy sector there are problems related to oil dependency, reliability, and CO₂ and NOₓ emissions. The transport system suffers from congestion, air pollution (particulates, NOₓ), energy use and CO₂ emissions. Cattle farming suffers from manure disposal problems, ammonia emissions and diseases like BSE and foot and mouth disease. These problems are deeply rooted in social production and consumption patterns.

Since the 1980s, much effort has been made to solve problems with product and process innovations. Cleaner products and processes have been developed alongside the application of end-of-pipe solutions. Sometimes these innovations have led to substantial improvements in environmental efficiency, such as in the case of automobile catalysts which greatly reduced tailpipe-emissions of pollutants. The focus in these cases has been on changing some technological artefact.

Substantial improvements in environmental efficiency (a ‘Factor 2’ is a general average) may still be possible with innovations of an ‘incremental’ kind. But larger jumps in environmental efficiency (possibly by a ‘Factor 10’) may only be possible with system innovations. The promise of transitions to sustainability via system innovations is schematically represented in Figure 1.1. Such transitions to sustainability require changes from, for example, one transport system to another or from one energy system to another. Such system innovations not only involve new technological artefacts, but also new markets, user practices, regulations, infrastructures and cultural meanings.

Because of its sustainability potential there is increasing interest from policy makers, NGOs and large firms in transitions and system innovations. The Stockholm Environment Institute, for instance, has published a book on the Great Transition (Raskin et al., 2002). The American National Research Council (1999) and the Dutch Research Council
(NWO) have made the study of transitions part of their research portfolio. The IHDP research programme (International Human Dimensions Programme on Global Environmental Change) has a Project on ‘Industrial Transformation’ (similar in meaning to ‘transitions’). The Dutch government gave transitions a central place in its fourth National Environmental Policy Plan (VROM, 2001) and has established ‘transition teams’ within various ministries.

To link with and to feed this growing interest, this book explores how system innovations come about and how policymakers might influence them.

**DELINEATING THE TOPIC OF ANALYSIS: TRANSITIONS AND SYSTEM INNOVATIONS**

In *Webster’s Dictionary* the term ‘transition’ is defined as a ‘passage from one state, stage, subject, or place to another’ or ‘a movement, development, or evolution from one form, stage, or style to another’. The states/forms have certain internal characteristics, which give them coherence and stability. The notion of a transition also has the connotation of rapid change, a ‘jump’ from one state to another.

Transitions can occur on different levels, depending on the unit of analysis. An example at the level of society as a whole is the transition from...
hunter-gatherer society to urban society. Another example is the transition from rural to industrial society. At a lower level, there are transitions in societal functions such as transport, communication, housing, feeding, energy supply and use, and recreation. Examples are the transition in transport systems from horse-and-carriage to automobile, or the transition from telegraph to telephone. There are also transitions at the level of organizations and firms, for example the transition from punched card machines to computers within IBM (Chandler, 2001) or the transition of DSM (Dutch State Mines) from coal mining via bulk chemicals to fine chemicals. This book focuses on a specific type of transition, notably transition at the level of societal functions.

What is it that changes at this level during transitions? In the way we use the term, these transitions involve changes in socio-technical systems. These comprise a cluster of elements, including technology, regulations, user practices and markets, cultural meanings, infrastructure, maintenance networks and supply networks. Technology plays an important role in fulfilling societal functions, but its functioning depends upon its relationship to the other elements. Technologies realize functionalities in concrete user contexts, which are made up of users, their competencies, preferences, cultural values and interpretations. User contexts are also shaped by a variety of existing artefacts and infrastructures (for example, road infrastructures, electricity networks), and regulations.

Technologies also need to be produced, distributed and ‘tuned’ with existing user contexts. This requires aspects such as technological knowledge, machines, skilled labour, capital, natural resources and components, and distribution networks. Although these supply and demand aspects can be distinguished analytically, they are mutually dependent in practice. To highlight this interrelatedness, we use the term ‘socio-technical systems’. Transitions at the societal level then involve a change from one socio-technical system to another, that is a system innovation.

This way of delineating the unit of analysis has several implications. Firstly, it means that the focus is wider than just an industry or a sectoral system of innovation. There has been some attention paid in the past to the emergence of new industries (Van de Ven and Garud, 1989), how industries coevolve with government and universities in so-called triple helix dynamics (Etzkowitz and Leydesdorff, 2000), and how firms, public authorities and universities work together in innovation systems or innovation communities (Breschi and Malerba, 1997; Malerba, 2002; Lynn et al., 1996). These approaches, however, mainly look at the supply-side and the production of innovations. They take the user side for granted or narrow it down to ‘the market’ which functions as a neutral selection environment.
There is another body of literature which shows that users do more than just buy and adopt (new) technologies. Cultural studies and social studies of technology have found that users have to ‘domesticate’ new technologies to fit existing user contexts. This involves symbolic and practical work, in which users integrate the artefact into their user practices, and cognitive work, which includes learning about the artefact (Lie and Sørensen, 1996; Du Gay et al., 1997). This fits an emerging trend in innovation studies and science and technology studies, in which more attention is paid to the role of users in innovation and technological development (see Schwartz-Cowan, 1987; Kline and Pinch, 1996; Eggerton, 1999; Coombs et al., 2001; Oudshoorn and Pinch, 2003). These observations imply that system innovations not only involve changes in industries, firms and technical knowledge, but also changes in user contexts and symbolic meanings.

This book acknowledges this and aims to bring together two bodies of literature which have remained relatively separate so far: on the one hand evolutionary economics, innovation studies and innovation system approaches and, on the other hand, cultural studies, and science and technology studies.

A second implication is that system innovations appear as a particular kind of innovation. To illustrate this, we use the innovation typology of Abernathy and Clark (1985), but widen it. They distinguish two dimensions. The first dimension consists of linkages between a firm and its customers, including channels of distribution and service, customer applications and customer knowledge. The second dimension relates to the technological and production competences of a firm, including production systems, skills, technical knowledge and supplier relations. Combining these two dimensions results in four types of innovations (see Figure 1.2):

1. architectural: disrupts existing technology and linkages with users;
2. niche creation: conserves existing technology but breaks linkages with users; new markets are explored with existing products;
3. incremental: conserves both existing technology and users;
4. revolutionary: disrupts technology but conserves user linkages (same markets).

Abernathy and Clark developed their innovation typology primarily to determine the consequences of different kinds of innovations for firms. While their point about linkages and alignments between elements is important, their focus on firms is too limited for our purposes as disruptions occur on a much wider scale during system innovations. Hence, system innovations
can be described as architectural innovations ‘writ large’, because they involve substantial changes on the supply side and on the user side. The term also highlights that system innovations are not about component changes, but about changes in the entire architecture or structure of socio-technical systems. Without changes on the user side, technological discontinuities are better described as ‘technological revolutions’, which do not change functionalities.

A third implication is that system innovations are multi-actor processes. This not only denotes interactions between actors within a societal group (for example, industry, user group, scientific community, policy community), but also interactions between societal groups. A range of societal groups or stakeholders is involved in system innovations: firms, suppliers, universities and knowledge institutes, public authorities, public interest groups, users. Their activities create and maintain elements of socio-technical systems. The societal groups have their own perceptions of the future, values and preferences, strategies, and resources (money, knowledge, contacts). Although these societal groups have some degree of autonomy, they are also related to each other and interpenetrate each other (Stankiewicz, 1992).

Their activities are to some degree aligned, and it is this that gives socio-technical systems their stability and a recognizable state or form. Within systems, innovations still take place but they are usually of an incremental nature, leading to trajectories in technical development, policies, infrastructures and demand. As long as these trajectories are aligned, socio-technical systems are stable. This stability is not the result of an overarching rationality or force by an all-powerful actor. Instead, stability is the emergent outcome of many activities of many actors. Stability need not be harmonious. There may be tensions and conflicts of opinion about a range of

Source: Abernathy and Clark, 1985.
matters, such as which problems should be highest on the problem agenda, which directions are most promising for solving a problem, or how resources should be allocated. When these tensions become pressing, a system may lose its stability, creating opportunities for change. But usually tensions remain manageable.

This discussion implies that, for our purposes, transitions or system innovations have the following main characteristics. First, they develop in a ‘coevolutionary’ way. They involve changes in both the supply side (in technology, knowledge, industry structures) and the demand side (user preferences, cultural meaning, infrastructure). Second, they are architectural innovations writ large, involving changes in the elements and structure of socio-technical systems. Third, they are multi-actor processes, involving a wide range of societal groups. A fourth characteristic that follows from a wide range of historical studies is that they unfold within a long timescale, possibly of the order of several decades (see Geels, 2002).

ACADEMIC RELEVANCE

Given the specifics of our interest in system innovation as described above, we identify three gaps in existing bodies of literature. The first gap concerns the ‘systems of innovation’ approach, which has emerged in the last decade. This approach investigates at different analytical levels how innovations emerge from the coevolution of a range of elements. The levels are national systems of innovation, regional systems of innovation or sectoral innovation systems. The main focus in the system of innovation approach is on the functioning of systems rather than the change of systems. For instance, at the national level these studies lead to a static or comparative analysis of the innovative performance of different countries. In a recent overview of sectoral systems of innovation, it was noted that ‘one of the key questions that need to be explored in-depth is: how do new sectoral systems emerge, and what is the link with the previous sectoral system?’ (Malerba, 2002; 262). This means that the topic of system innovation as we use it is under-addressed in this literature.

A second gap stems from the literature on ‘path dependence’ and ‘lock-in’. In evolutionary economics, David (1985) and Arthur (1988) have shown that path dependence plays an important role in the case of two competing technologies. Once one of the technologies has gained a lead, it benefits from increasing returns to adoption and creates a dominant path. Several mechanisms cause increasing returns, such as economies of scale, leading to lower cost, learning-by-using, network externalities,
informational increasing returns, and technological interrelatedness. Because of these increasing returns a certain technology becomes entrenched while there is no guarantee it is the ‘best’ one from a broader societal perspective. Other economists have widened the analysis, adding institutional aspects and user routines to the lock-in analysis (Cowan, 1990; Cowan and Gunby, 1996).

Research from other disciplines has added more reasons why existing systems are characterized by stability, inertia and lock-in. Established systems may be stabilized by legally binding contracts (Walker, 2000). Actors and organizations are embedded in interdependent networks (with suppliers or users), which represent a kind of ‘organizational capital’, and create stability through mutual role expectations. Cognitive routines make engineers and designers look in particular directions and not in others (Nelson and Winter, 1982; Dosi, 1982). This can make them ‘blind’ to developments outside their focus. Core capabilities can thus turn into core rigidities (Leonard-Barton, 1995). Firms have sunk investments and built-up capital, which they do not want to write-off (examples are investments in machines and production tools, skills and knowledge). It is difficult for established firms to switch to competence-destroying breakthroughs (Tushman and Anderson, 1986; Christensen, 1997).

Existing systems are also stable because they are embedded in society. People adapt their lifestyles to them, favourable institutional arrangements and formal regulations are created, and accompanying infrastructures are set up. The alignment between these heterogeneous elements leads to ‘technological momentum’ (Hughes, 1994). The importance of these alignments between heterogeneous elements is highlighted in such concepts as the ‘techno-institutional complex’ (Unruh, 2000) and ‘techno-economic networks’ (Callon, 1991). All these approaches highlight aspects of the stability of existing systems but none of them addresses the issue of change and transition from one system to another. Given all these explanations of stability, it is a mystery how and why transitions occur. Path-dependency literatures may help us understand lock-in, but how can we understand ‘lock-out’?

A third gap relates to recent academic sustainability debates. There has been a widening in recent years of the analytical focus, from clean products to sustainable systems (Schot et al., 1994; Vellinga and Herb, 1999; Unruh, 2000; Jacobsson and Johnson, 2000; Berkhout, 2002). In transport, energy and other systems there are promising new technologies with better environmental performance. But many of these new technologies are not (yet) taken up. This is partly for economic reasons, but there are also social, cultural, infrastructural and regulatory reasons. Existing systems seem to be
‘locked in’ on many dimensions. Implementation of promising new environmental technologies may require other changes in user practices, regulation or infrastructure. Although the importance of system innovations is increasingly emphasized in sustainability debates, there is not yet much known about how system innovations occur and how policy makers may influence them.

GOVERNANCE

Transitions are complex, uncertain and involve multiple societal groups or stakeholders. Hence, policymakers and other decision-makers puzzle over how to influence system innovations and how to identify possible and promising directions for transitions. This is complicated by the awareness that the state is not an all-powerful and all-knowing actor in this matter. Public authorities are just one societal group among several others. Like other groups, they have limited power, a limited cognitive perspective and limited resources to influence system dynamics.

This observation caused a shift in policy studies from a focus on ‘government’ to ‘governance’ (e.g. Kooiman, 1993; Rhodes, 1997; Kohler-Koch and Eising, 2000; Van Heffen et al., 2000). Governance means that there is directionality and coordination at the systems level, but that it has an emergent character, arising from the interaction between multiple societal groups. Public authorities may try to influence this emergent directionality, but cannot steer it at will. This emerging governance paradigm emphasizes aspects such as policy networks, interaction between multiple societal groups and learning processes.

This is not the only relevant policy paradigm. In policy science, three general policy paradigms are distinguished (see Table 1.1): (i) the traditional top-down model with a central role for (national) government and hierarchical relations, (ii) a bottom-up or market model with a large degree of autonomy for local actors, and (iii) a governance or policy network model with shared rule-making and agreements between interdependent actors with diverging values and beliefs. These three policy paradigms differ not only in their basic philosophy, but also in their instruments. Formal rules and regulations are common in the command-and-control paradigm, subsidies, taxes and (financial) incentives in the market model, and network management, learning processes, experiments and interactive policymaking in the third paradigm.

These policy paradigms coexist in all democratic societies with varying degrees of emphasis on each of them. This variety of paradigms and instruments complicates the issue of governing transitions and system
innovations. It raises questions such as these: is one policy paradigm best suited to influence transitions or is a mix of paradigms and instruments needed? In the latter case, what should this mix look like? Is the optimal mix dependent upon specific circumstances and, if so, which ones?

<table>
<thead>
<tr>
<th>Table 1.1 Different policy paradigms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level of analysis</strong></td>
</tr>
<tr>
<td><strong>Perspective</strong></td>
</tr>
<tr>
<td><strong>Characterization of relationships</strong></td>
</tr>
<tr>
<td><strong>Characterization of interaction processes</strong></td>
</tr>
<tr>
<td><strong>Foundation scientific disciplines</strong></td>
</tr>
<tr>
<td><strong>Governance instruments</strong></td>
</tr>
</tbody>
</table>

*Source: Based on De Bruijn et al., 1993: 22.*
RESEARCH QUESTIONS AND LEVEL OF AMBITION

Given our dual ambition to enhance the understanding of transitions and to stimulate the formulation of policies that guide transitions towards sustainability there are two main research questions that drive this book:

1. How do system innovations or transitions come about? What theories can be used to conceptualize (part of) their dynamics and what gaps exist in those theories? What can we learn from historical examples of transitions?
2. How can transitions or system innovations be influenced by actors, in particular by public authorities? What instruments and tools exist and how should they be used?

Although this book aims high, it is not our ambition to provide the ultimate answers to these questions. Instead, we want to create a signpost into this uncharted territory. System innovations are a complex topic, involving many kinds of actors and issues. In this book, authors from different disciplines make their own distinctive interventions into the topic, coming at it from different angles and with different intellectual frameworks: innovation studies, sociology of technology, institutional economics, history of technology, policy studies, including studies of network governance, learning and the impact of regulation, innovation management and governance approaches, organizational studies and management of structural change and leadership. This grouping of different disciplinary backgrounds around a particular topic creates variety and space for interdisciplinary discussion. Although different disciplines highlight different aspects of system innovations, they do share a common view on human actors as boundedly rational and embedded in social networks and institutions. This means transitions are not and cannot be planned in advance in a rational manner but emerge as actors navigate their way through multiple uncertainties. These shared views provide common ground between different authors.

BOOK STRUCTURE

The book is in two sections, addressing the two main research questions. The first section is on ‘Understanding Transitions’, with one part focusing on theoretical explorations (Chapters 2, 3 and 4), and one on empirical examples (Chapters 5 and 6). The second section deals with ‘Inducing Transitions’; it also consists of two parts, the first on transition management.
in general (Chapters 7 and 8) and the second on tools for transition man-
agement (Chapters 9, 10 and 11).

**Part I: Theoretical Explorations of Transitions**

Chapter 2 by Frank Geels addresses the general question of how system innovations come about. Geels reviews a broad range of relevant litera-
tures, concluding that they do not add up to an integrated perspective on system innovations. He offers a pragmatic synthesis in the form of a ‘multi level perspective’ (MLP) to analyse and explain transition processes. These levels are (i) technological niches where novelties are devel-
oped, (ii) socio-technical regimes and (iii) socio-technical landscape, which comprises a range of exogenous developments which influence regimes and niches. The main argument is that system innovations result from link-
ages between processes at these multiple levels. This means that system innovations are not caused by a change in a single factor or ‘driver’, but are the result of the interplay of many processes and activities.

Chapter 3 by Frans Berkhout, Adrian Smith and Andy Stirling also looks at the general level of system innovations. They take aim at the multi-
level perspective, arguing that this approach places too much emphasis on the role of technological niches as the principal locus for regime change. Instead, they argue, there is a range of different transition contexts in which regime change can take place. They argue that there is a greater plurality of possible transformation pathways than suggested by the multi-level per-
spective notion. They develop a four-fold typology of transition contexts, which they illustrate with brief examples.

Chapter 4 by Elizabeth Shove has a more specific focus on users and consumption. She sees much of the current work on transitions as socio-technical in its orientation as it acknowledges the institutional and political processes required in support. She argues, however, that the agenda remains lopsided, skewed around provision rather than consumption and around the diffusion rather than the use of technological systems, tools and techniques. She seeks to recover some of that missing ground. Using the case of laundering, she argues that it is necessary to think more systematically about the relation between consumption, provision and prac-
tice. She suggests that shared understandings of ‘normality’ are import-
ant in this respect. Notions of what it is to be a normal and acceptable member of society have far reaching environmental implications. They carry in their wake a trail of resource requirements like those associated with daily showering, with wearing freshly laundered clothing, with not having a siesta, with eating imported food or with having foreign holidays.
Part II: Empirical Examples of Transitions

Chapter 5 by Frank Belz presents a historical analysis of changes in the Swiss agri-food chain over the past three decades. He describes a shift away from the industrialized form of agriculture, a form which creates major sustainability problems. The shift is not yet completed but has progressed a long way. Switzerland is one of the leading Western countries in sustainable agriculture, balancing economic, ecological and social dimensions. In this transition two new forms of agriculture play a role. Organic farming takes a holistic stance, respecting the principles of nature, by seeking to maintain long-term fertility and biological activity of soils using locally adapted biological and mechanical methods as opposed to reliance on external inputs. Integrated production is a third way between organic farming and industrialized agriculture. In his analysis, Belz proposes a number of additions to Geels’s multi-level perspective.

In Chapter 6, Aad Correljé and Geert Verbong describe the transition to the use of natural gas in the Netherlands in the 1950s and 1960s. The discovery of a large deposit of natural gas in 1959 caused a shock to the energy system based on coal and coal-based gas. They distinguish three dimensions of the gas regime, notably (i) the material network, (ii) the institutional framework and (iii) the market for energy. They show that the transition to a new system required a process of interrelated changes on these three dimensions. They also show the transition cannot be understood without taking note of earlier developments that took place in the 1950s. The authors analyse in detail the strategies, visions and activities of relevant actors, and show the struggles and negotiations that took place.

Part III: Transition Policy

In Chapter 7, René Kemp and Jan Rotmans present a general framework for transition management. They argue that policy interventions should target not just economic conditions (through taxes and regulations) but also beliefs, expectations and institutional factors. They propose a management strategy based on modulation of ongoing dynamics rather than planning and control. The overall steering philosophy is to embark on a process of ‘learning-by-doing’. This involves articulation of future visions, setting up experiments to learn about the feasibility of visions, and the evaluation and adjustments of visions. Transition management is not a one-off exercise but involves several policy cycles of adjustment and learning. In that sense it is goal-oriented incrementalism. The authors apply their ideas to the domain of transport and mobility. They show what the first step of transition management might look like for a transition towards sustainable transport.
In Chapter 8 Geert Teisman and Jurian Edelenbos argue that ‘management of transitions requires a transition of management’, that is, a transition from hierarchical steering to interactive forms of governance. This requires institutional change because, as they state it, ‘the best way to kill a new idea is to put it in an old organization’. The authors identify three barriers which hinder transition management: (i) missing links between interactive processes and formal decision-making; (ii) fragmented departmental structures of governmental organizations frustrating productive and innovative interactions; and (iii) the reluctance of public actors to share responsibility and accountability with each other and with private actors or societal actors. They discuss several experiments as possible forerunners for new democratic governance systems: parallel democracy, hybrid democracy and participatory democracy. In current Dutch practice, the first form is advocated as a model for system innovation. The authors argue that this is not sufficient and that a change should take place towards a hybrid democracy or participatory democracy. This requires a redefinition of the role of various actors, especially at various levels of government.

Part IV: Tools for Transition Policy and Empirical Illustrations

In Chapter 9 Halina Brown, Philip Vergragt, Ken Green and Luca Berchicci discuss ‘bounded socio-technical experiments’ (BSTE) as attempts to introduce a new technology, service, or social arrangement on a small scale. Based on insights from theories of organizational learning, policy-oriented learning and diffusion of innovation, the authors identify two types of learning: technical single-loop learning, and higher-order social learning. The first type of learning occurs among the participants in the experiment and their immediate professional networks. The second type occurs in society at large. The authors argue that both types play a key role in a transition towards sustainable mobility systems. They analyse two Dutch experiments in personal mobility, the development of a three-wheeled ‘bike-plus’ vehicle called Mitka and an attempt to solve mobility problems on the island of Texel. The cases show that the first type of learning took place to a considerable extent and that it can be facilitated by deployment of structured visioning exercises, by diffusion of ideas among related BSTEs, by innovative couplings of problems and solutions, and by creating links among related experiments. The cases also show that the second type of learning was more difficult. The authors provide recommendations on how both kinds of learning could be organized, stressing the importance of visions and vision-building processes.

In Chapter 10 Sirkku Kivisaari, Raimo Lovio and Erja Väyrynen take as a starting point that experimenting with alternatives to an existing system...
can play a crucial role in broader transition processes because they provide the seeds for change. They use the so-called ‘societal embedding of innovations’ approach to analyse management of experiments. This approach has been designed to enhance commercialization of innovations that yield financial profit as well as contribute to sustainable development. It has been geared especially towards supporting collaboration between public and private actors in cases where there is a considerable public interest in finding innovative solutions to societal issues. The chapter discusses two Finnish experiments, which can be perceived as pilots of system innovations. The first deals with a new energy service company concept in Finnish municipalities and the second with development of a new diabetes self-management system. Combining findings from both cases, they discuss how the management of these experiments can be strengthened so that their results can indeed form the seeds for a transition.

In Chapter 11 Boelie Elzen, Frank Geels, Peter Hofman and Ken Green present a new scenario method to explore future system innovations and support transition management. In a strict sense, transitions cannot be steered, because of their complex nature, but it is possible to stimulate developments in more sustainable directions over a longer period of time. This requires a vision of which directions that might be, that is which combination(s) of technologies and their societal embedding might contribute to a sustainable system. To help develop such visions, scenario studies or other foresight methods can be used. Although many such methods exist, the authors argue that they have limitations for exploring system innovations and transitions. Hence, they present a new scenario method, called ‘Socio-Technical Scenarios’. The chapter describes the main features of the method and illustrates it by describing two short scenarios for the passenger mobility domain. The authors thus provide a concrete tool to help develop guiding visions.

In Chapter 12 the editors take stock of the findings of the book, and suggest a research agenda for the future.

REFERENCES


VROM (2001), *Een wereld en een wil, werken aan duurzaamheid, Nationaal Milieubeleidsplan 4*, Dutch Environmental Policy Plan 4, established by nine Dutch departments under the coordination of VROM (The Department of Housing, Spatial Planning and Environment), The Hague: VROM.
