Science, Technology & Human Values
http://sth.sagepub.com

Feelings of Discontent and the Promise of Middle Range Theory for STS: Examples from Technology Dynamics

Frank W. Geels

Science Technology Human Values 2007; 32; 627 originally published online Sep 12, 2007;
DOI: 10.1177/0162243907303597

Science, Technology, & Human Values
Volume 32 Number 6
November 2007 627-651

10.1177/016224390303597
http://sth.sagepub.com
hosted at
http://online.sagepub.com

Feelings of Discontent and the Promise of Middle Range Theory for STS

Examples from Technology Dynamics

Frank W. Geels
Eindhoven University of Technology

This article critically discusses the state of STS, expressing feelings of discontent regarding four aspects: policy relevance, conceptual language, too much focus on complexity, theoretical styles. Middle range theory is proposed as an alternative, promising avenue. Middle range theories focus on delimited topics, make explicit efforts to combine concepts, and search for abstracted patterns and explanatory mechanisms. The article presents achievements in that direction for technology dynamics, particularly with regard to the role of expectations, niche theory and radical innovation, and the multi-level perspective on sociotechnical transitions.

Keywords: *technology dynamics; mechanisms; expectations; emerging trajectories; sociotechnical transitions*

Introduction

Talk of 'middle range theory' (MRT) is often an indication of discontent in a discipline, suggesting a middle way between undesirable extremes.[1] Hence, the term is difficult to define and may change content between fields and over time. In this section, I first draw on debates in sociology and organization theory in order to explicate what I mean by MRT. I then identify four problems with science and technology studies (STS) and suggest how MRT might form a solution to these problems. In section 2, I illustrate how and where substantial achievements in middle range theorizing have already been made in the area of technology dynamics. Section 3 offers some concluding remarks about how the examples in section 2 provide solutions to the problems described in section 1.

Merton introduced the notion of MRT in sociology in the three editions of *Social Theory and Social Structure* (1949, 1957, 1968). The core of his concern was that the development of an all-encompassing system of concepts would prove both futile and sterile. In criticizing Parson's structural-functionalism, Merton (1948) noted that "To concentrate solely on the master conceptual scheme for deriving all sociological theory is to run the risk of producing twentieth-century equivalents of the large philosophical systems of the past, with all their suggestiveness, all their architectonic splendor and all their scientific sterility." Merton not only attacked grand theoretical systems, but was also critical of the opposite strategy of constructing inventories of low-level empirical propositions. C. Wright Mills (1959) also complained about a "general malaise of contemporary intellectual life" (p. 19), diagnosing that sociology was increasingly divided between grand theory, which addressed a "level of thinking so general that its practitioners cannot logically get down to observation" (p. 33), and abstracted empiricism, which focused on data-collection and statistical runs to find correlations between variables.

Merton proposed MRT as a middle way between both extremes, defining them as "theories that lie between the minor but necessary working hypotheses that evolve in abundance during day-to-day research and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization and social change" (Merton 1968, 39). This implies that MRT relates to theoretical *scope*. Empirical generalizations are located at the small-scope end of the continuum, while grand theories are located at the large-scope end. Middle range theories are located between the two extremes. Merton gave a few examples: self-fulfilling prophecies, reference group theory, role-sets, dissonance theory and the two-step flow model. Middle range theories such as these do not consist of elaborate frameworks with endless conceptual distinctions. Instead, MRT consists of a limited set of interrelated propositions, aimed at understanding limited topics. MRT is not about broad, abstract entities such as 'society' or 'social system'. Instead, MRT addresses delimited aspects of social phenomena. Another characteristic is that concepts and propositions should be clear, specific and empirically researchable. MRT differs from grand theory, because it emphasizes interaction between theory and empirical research. A third characteristic is that middle range theories specify relationships between concepts into analytical models: "An array of concepts does not constitute theory. . . . It is only when such concepts are interrelated in the form of a scheme that a theory begins to emerge. Concepts then constitute the definitions or prescriptions of what is to be observed.

They are the variables between which empirical relationships are to be sought" (Merton 1968, 143). Analytical models and mechanisms are not deterministic, but indicate how one concept influences another.

In organization theory, Pinder and Moore (1980) proposed MRT in an edited volume containing 35 chapters from different contributors. In the introduction, the editors mentioned several problems that plagued the field. The first was an overemphasis on methodological rigor and hypothesis *testing* at the expense of hypothesis *construction*. "We have become preoccupied with measurement precision, with the creation and internal validation of new instruments" (p. 7). Much work was seen as narrowly based, extending existing hypotheses and relying on previously accepted assumptions. Another problem was disappointment over the results of present research efforts. The gains in knowledge about how organizations should be structured, how leaders should lead and motivate people, how jobs should be designed, was seen as modest. Possible explanations for low explanations of variance were faulty measurement, poor sample selection, poor theory, and poor hypothesis construction. A third problem was the loss of parsimony. "Our theoretical language seems full of fuzzy concepts and of similar but not identical definitions of terms, resulting in an inability to achieve parsimony. ... We should be concerned with standardization in defining the variables we include in our hypotheses" (p. 9).

Because of these problems, the editors felt that the dominant mode of theory-building required modification. They did not want to return to grand theory, which had dominated between 1920 and 1950.[2] These grand theories were stated at a level of abstraction such that it was impossible to deduce hypotheses amenable to disproof. MRT was proposed to avoid the pitfalls of grand theory and solve the problems in present theory building. It was especially hoped that MRT would stimulate creativity: "What must be sought is an approach to theory building that ensures adequate emphasis on creative idea and hypothesis generation" (p. 7). Despite its 35 chapters, however, the book did not clearly demarcate what MRT was and how it should be practiced.

This is a general problem. Because MRT is usually an expression of discontent, it is clearer what it should *not* be, than what it *is*.[3] Also Boudon (1991), in an attempt to explain MRT, did so primarily in a negative sense. MRT means "it is hopeless and quixotic to try and determine the overarching independent variable that would operate in all social processes . . . or to find out the two, three or four concepts . . . that would be sufficient to analyze all social phenomena" (p. 519). Boudon further discusses some examples (reference group theory, two-step flow model, subjective rationality theory), and gives the following general delineation of MRT: "a set of statements that organize a set of hypotheses and relate them to segregated observations. If a 'theory' is valid, it 'explains' and in other words 'consolidates' and federates empirical regularities which on their side would otherwise appear segregated. This amounts also to saying that mere empiricism is of little worth" (p. 520).

Given this characterization of MRT, we can ask which feelings of unease underlie this special issue about MRT in STS? In my view, there are four important aspects.

Limited Impact on Policy and Other Disciplines

Reflecting on the past 25 years, Edge (2003) praised STS as a vibrant and dynamic academic field, but also expressed concerns about its limited impact on policy and other disciplines. I share his diagnosis, and see three explanations. First is the proliferation of concepts and jargon, which are hard to understand for outsiders. Second is the emphasis on complexity, local situatedness and contingency, which makes it difficult to communicate general lessons. And third is the normative aversion against instrumental contributions, for fear of technocracy. Hence, most policy contributions are critical or reflexive, criticizing policy discourses for their 'wrong' assumptions about technology and society. Popular recommendations also concern the empowerment of outsiders and lay persons and calls for more public participation. With regard to public participation, I agree with Collins and Evans (2002) who feel "a little uncomfortable when every treatment has the same political recipe, because it makes it all too easy to imagine that the prime motivation is political rather than analytical" (p. 263).

MRT cannot solve all these problems. But the articulation of medium-range patterns and stylized conceptual models may enhance STS's policy relevance, especially if they address issues that policy makers struggle with.[4]

Concepts Are Not (Yet) Theory

The proliferation of concepts not only hinders communication with outside audiences, it is also a particular theoretical style. Too many articles only introduce another concept, and illustrate it with a case study. But loose concepts are not yet theory (although they are part of it). Furthermore, concepts are often fuzzy, undefined and not demarcated with regard to similar concepts. Molina (1995, 387) criticizes STS concepts such as heterogeneous engineering, enrollment, system building, because "most of the meanings have yet to establish a set of conceptual tools that makes them analytically operational beyond the metaphor". More importantly, there are few efforts to systematically combine and interrelate concepts. Hence Molina (1995, 387) argues that "A need continues to exist for further refining and developing systematic approaches to understand complex technological processes by integrating concepts in an analytically operational way". In this respect, MRT may make useful contributions to STS.

Too Much Theoretical Emphasis on Complexity, Local Practices and Contingency

Taken-for-granted ideas about the linear model and technological determinism were important 'enemies' in the early days of STS. Social construction of technology (SCOT) and actor-network theory (ANT) successfully undermined the intellectual basis of these ideas. In detailed case studies of artifacts and local projects, SCOT and ANT 'followed the actors' and showed the complexity and messiness of interactions between actors, e.g., disagreements, controversies, negotiations, battles, coalitions. The approaches emphasized contingency and non-linearity, implying that 'things could have been different'. Although these perspectives and the many case studies served their purpose well in the 1980s and 1990s, they are now suffering from diminishing returns. There is too much repetition of the same message. According to Guggenheim and Nowotny (2003), this repetition stems from the role STS gave itself. An implicit assumption of the early program was that the prevailing understanding of science and technology was flawed, and that it was the mission of STS to correct these distortions. STS took on the role of Modern Sisyphus, whose task was not only hard, but also repetitive. Because scientists and 'society' continue to make 'wrong' claims about science and technology, STS's task is never-ending. But claims about complexity, contingency, local situatedness, and messiness may become repetitious and unsurprising.

MRT suggests a new theoretical avenue, aimed at finding patterns, regularities and stylized mechanisms. This avenue accepts and cherishes the importance of complexity, messiness and local practices, but maintains that (inter)actions in many local practices can add up to patterns and regularities at a global level. The task of MRT is thus to stylize and simplify complexity.

Theoretical Styles

Theory is not only about concepts and propositions, but also about style. DiMaggio (1995) usefully distinguished three styles. The first is 'theory as covering laws' (variance theory). In this style the world is approached as made up of variables; the aim is to explain variance rather than regularities.

Theory attempts to cover 'laws' and correlations between variables. Although this theoretical style is hardly practiced by STS (for good reasons), it is the bread and butter of economists, psychologists and many mainstream sociologists. The second style is 'theory as enlightenment', which aims to clear away conventional notions and assumptions to make room for artful and exciting new perspectives. Theory may be defamiliarizing, rich in paradox, deconstructing accepted assumptions. ANT, SCOT and reflexivity approaches have characteristics of this style. Law (1999, 3), for instance, praises ANT for its "shock value" and "potential for scandal", because "sacred divisions and distinctions have been tossed in the flames". And Woolgar (2004) values the potential for provocation and deconstruction, which is supposed to enable people to remove their blinkers. DiMaggio's third style is 'theory as narrative' (process theory). These theories give an account of a social process, allowing entities (actors, structures) to change their identity during the process, which is seen as interpreted and enacted. These theories care less about the variance explained, and more about empirical tests of the plausibility of the narrative. It requires that hypotheses detailing regularities in relations among variables be accompanied by plausible accounts of how the actions of real humans could produce the associations predicted and observed. Clearly, much STS work also belongs to this style.

These styles differ with regard to criteria for 'good theory'. Weick (1999) proposes that all theories make tradeoffs between three such criteria: (a) generality/scope; (b) simplicity/parsimony (Ockham's razor), and (c) accuracy/specificity. A theory that satisfies two characteristics is less able to satisfy the third characteristic. 'Theories as covering laws' are usually strong in *generality* and *parsimony*, but weak in *accuracy* (e.g., economic models based on rational-choice assumptions). 'Theories as narrative' are somewhat more complicated. Tradeoffs between *accuracy* and *simplicity* are common in ethnography, narratives, cases and idiographic inquiry. SCOT, however, is *general* and *simple*, but not very *accurate/specific*. The conceptual framework is *simple*, because it uses only a handful of concepts (interpretative flexibility, relevant social groups, closure, stabilization) which are loosely related. *Complexity* occurs primarily in the empirical operationalization of these concepts in case studies, which requires elaboration and detail. SCOT is *general*, because it is about the whole process of technological change, and because the concepts are abstract and can apply to many social domains (e.g., politics, culture). SCOT's conceptual framework is not very *detailed/specific*, but more a sensitizing scheme for case studies. In this sense SCOT is not a 'grand theory'. Furthermore, concepts such as 'closure' are relatively opaque and the underlying mechanisms are unclear. While Bijker (1995) emphasizes closure through negotiation, others suggested conflict and power (Hård, 1993) or market processes. 'Theories as enlightenment' are more difficult, and may criticize the three criteria for 'good' theory. Actor-network theorists, for instance, doubt the importance of *simplicity*.[5] They value other criteria more highly, e.g., creativity, revelatory power, appeal, and surprise. Furthermore, ANT is difficult to point down, because it not only addresses dynamics of science and technology, but also forms a complex ontology. Its theory of dynamics is relatively *simple*, because it consists of a few main concepts (translation, delegation, enrollment, association, alliance, obligatory passage point, actant, negotiation, immutable mobiles, inscriptions). But the ontology is quite *complex*, undermining established cognitive conventions[6] and developing alternative vocabularies. Furthermore, ANT is *general*. Its concepts not only cover science and technology, but can be applied to all domains of human life. ANT is not a grand theory in terms of elaborateness and specificity, but it does have grand ambitions. It wants to transform social science by inventing a "symmetric metalanguage" (Callon and Latour 1992, 354) and "a common vocabulary and a common ontology" (p. 354) that "crisscross the divide of two asymmetric vocabularies" (p. 359). But ANT is not very *accurate/specific*. Its concepts are descriptive and idiosyncratic, something for which Latour (1999) praises ANT.[7] He characterizes ANT not as a theory, but as "a method . . . to systematically record the world-building abilities of the sites to be documented and registered" (p. 20-21). This makes ANT vulnerable to Collins and Yearley's (1992) charge of empiricism. They claim that ANT is strong at description (following the actors), but weak at explanation. Callon and Latour (1992: 362) seem to agree with this diagnosis, arguing that ". . . explanations might not be desirable after all . . . A complete description of network dynamics might provide a better explanation, in the end, than the delusive search for causes". This means that explanation becomes almost the same as description. "Actor-network theory almost always approaches its tasks empirically" (Law 1992, 185). Elsewhere, Law (1996) suggests that general patterns may not exist, and that all we can do is tell "lots of little stories".[8]

I conclude from this discussion that two important STS theories (SCOT and ANT) are characterized by a gap between relatively simple, sensitizing conceptual schemes and detailed, complex case descriptions with some empirical generalizations. This is not exactly similar to Merton's gap between grand theory and abstracted empiricism. The theories are not elaborate and grand (although ontological ambitions are) and empirical data are not abstracted (but detailed and full of life). It is a gap nevertheless. In a review of theories of technological change, Rip and Kemp (1998) diagnosed: "The new history and sociology of technology has shed light on forgotten views and failed technological directions, but it has not progressed very far in the direction of theory. … At this stage, explanations of the eventual shape of technology tend to be glosses on specific case studies, informed by general sociological theories" (p. 358-359).

Some scholars suggest that this problem plagues sociology more widely: "It turns out that what often goes under the rubric of social theory, should more properly be viewed as conceptual or sensitizing schemes, and not as explanatory theory proper. Much of modern social theory has a tendency … to label, relabel, and to describe rather than to explain" (Hedström and Swedberg 1998, 1). Sociology finds it hard to abstract because "unlike the humanities and economics, sociology places high value on empirical truth, perhaps too high a value" (Davis 1994, 182). There is a strong belief in explanations that provide accounts of what happens as it actually happens. For STS, Bijker (1993, 117) also warned that "our endeavor" should not turn "into mere storytelling and thereby into a far less ambitious project". He argued that the ideal of explanation should not be given up and that some forms of reduction are necessary.

In response to this problem Hedström and Swedberg (1998) advocate the search for medium-scope patterns and mechanisms. They distinguish between a complex social reality and an intentionally simplified analytical model of this reality. Such analytical models provide explanations of demarcated topics with a limited number of interrelated concepts. The patterns and mechanisms derive from abstractions and analytical accentuation, which always distort the descriptive account of what actually happened. "The standard sociological critique of analytical theory focuses on the reality of its assumptions. … But criticizing an analytical model for lack of realism is a common sense instance of the logical fallacy, which consists of mistaking the abstract for the concrete. … All models by their very nature distort the reality they are intended to describe. The choice must be guided by how useful the various analytical models are for particular purposes" (Hedström and Swedberg 1998, 14-15).

I think this suggestion, which may be characterized as MRT, is also fruitful for STS. This does not mean that STS needs to practice 'covering laws'. Mechanisms are not the same as laws. Mechanisms are based on actions not correlations between variables. It is actors and not variables who do the acting. The search for patterns and mechanisms means sacrificing some of the empirical accuracy and complexity in order to gain more abstracted, general MRT about delimited topics.

Middle Range Theory in STS

Given this explication of aspects of discontent, MRT in STS would have the following characteristics. It would not address the whole of science and technology, but focus on limited themes and topics. It would make explicit efforts to combine different concepts in an analytical model and it would search for patterns and explanatory mechanisms.

MRT aims to do justice to all three criteria of 'good theory' by trading off a little bit on each of them. On *generality and scope*, MRT is between empirical generalizations and theories that address the whole of science and technology. On *simplicity*, MRT does not consist of complicated, elaborate conceptual frameworks, nor of general, fuzzy concepts; instead, it is about a limited number of related concepts. On *accuracy*, MRT is about propositions with a clear link to empirical cases. But the patterns and mechanisms are abstract, and accept some loss of empirical complexity.

Middle range theorizing in STS is not just a wish or empty program. Substantial achievements have already been made, although these are not always visible, because they are published in book chapters or journals outside the 'STS core'. Examples of such achievements are the subject of the next section.

STS Achievements in Middle Range Theory

This section discusses three examples of middle range STS theories in technology dynamics. The examples are based on multiple case studies by myself and others on the role of expectations, on niche theory and on sociotechnical transitions.

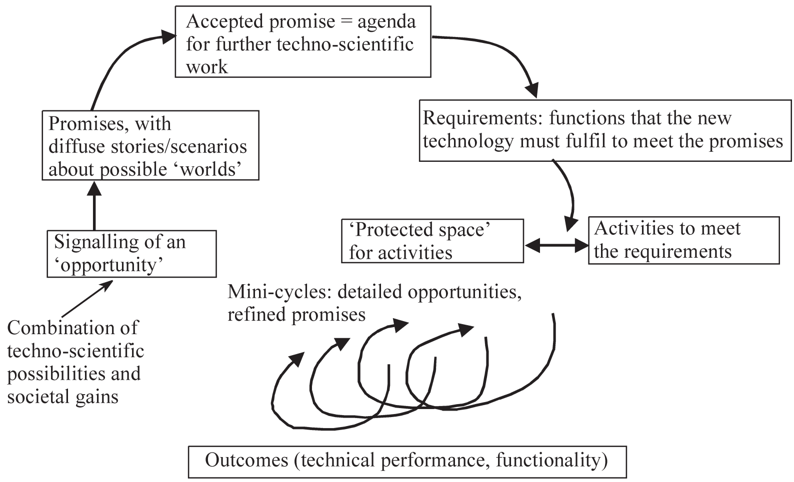
The Role of Expectations and Promises in Technology Dynamics

Several scholars have addressed the role of expectations in technological change (Van Lente and Rip, 1998; Brown and Michael, 2003; special issue in *Technology Analysis & Strategic Management*, 2006, No. 2-3). Expectations can act as self-fulfilling prophecies, because they guide social actions in technological change. This is an MRT, because it focuses on a particular topic, specifies relationships between concepts and has clear mechanisms.

When new technologies emerge there is much uncertainty about technical characteristics and user context. There is no guarantee of success. Actors invest in the development of novelties, because they *expect* them to have bright futures. Initial expectations have the form of 'diffuse scenarios', sketching a future world in which the product can be sold. These diffuse scenarios involve assumptions about users, markets, regulations, and technical progress (see also Callon's (1986) 'actor-worlds'). Product champions use expectations strategically. They make promises about future worlds, profits and societal benefits to attract attention and resources from managers, policy makers, and other sponsors. If these promises are accepted and stabilize into an agenda, they begin to exert force on further activities. Expectations are then translated into requirements, indicating directions for R&D activities. This creates a 'protected space' for technology-development actors, who receive resources to make the expectations come true. When projects end, usually after a couple of years, outcomes are assessed and new promises are formulated. When expectations are not yet realized, 'repair work' is performed. Actors give reasons for delay and make new promises that are more specific, refined and detailed. This results in 'promise-requirement cycles' (Van Lente and Rip, 1998) that can be continued as long as there is sufficient progress to legitimate new promises and as long as sponsors continue to believe them (Figure 1). Through subsequent cycles a shift occurs from broad, diffuse expectations to more concrete ideas about technical form and social function. Because expectations guide R&D activities that work towards realizing them, they act as self-fulfilling prophecies.

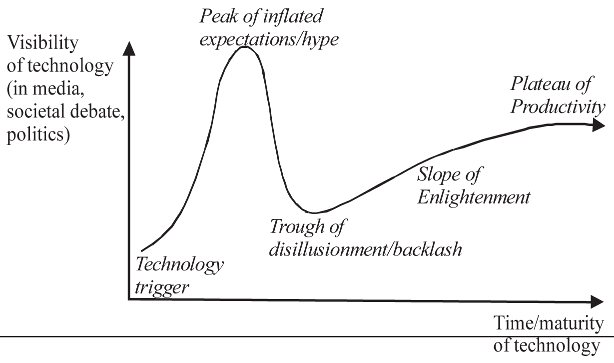
The promise-requirement cycle is a mechanism that explains patterns such as the hype-cycle. According to Gartner a hype-cycle consists of five phases (http://www.gartner.com/pages/story.php.id.8795.s.8.jsp): (a) Technology trigger: a public demonstration, product launch or other event generates significant press and industry interest; (b) Peak of inflated expectations: a flurry of well-publicized activity leads to over-enthusiasm and unrealistic expectations; (c) Trough of disillusionment: technical setbacks occur, leading to sentiments that the technology does not live up to its inflated expectations; it rapidly becomes unfashionable, and the press abandons the topic for the next hot thing; (d) Slope of Enlightenment: gradual improvement enhances understanding of the technology's applicability, risks and benefits. When real-world benefits are demonstrated, the new technology may gradually diffuse; (e) Plateau of Productivity: the benefits become widely demonstrated and accepted; the technology becomes increasingly stable and evolves in second and third generations (Figure 2).

Figure 1
The Promise-Requirement Cycle in Technological Development



Source: Van Lente and Rip (1998, 223); Geels and Smit (2000, 881).

Figure 2
The Hype-cycle



Source: Adapted from Flen and Linden (2005) (http://www.gartner.com/DisplayDocument?id=484424; accessed March 30, 2007).

Although this pattern does not necessarily have the happy ending Gartner suggests, the first three stages are quite common. The pattern's explanation is that outcomes and technical learning processes become uncoupled from expectation dynamics. The discussion of promising new technologies in the public domain may acquire a dynamic of its own. Because technical development activities are sheltered in 'protected spaces', the wider world may not know of the setbacks, technical bottlenecks and problems. The stock market may act as a further mechanism to boost expectations, leading to a financial bubble (identified retrospectively). As stocks of the new technology rise, expectations become rosier, more people will buy them, and stocks rise further (Perez, 2002). When problems and difficulties in developing the new technology eventually become known, high expectations may suddenly turn sour, leading to a backlash and stock market collapse.

Niche Theory and Emerging Technical Trajectories

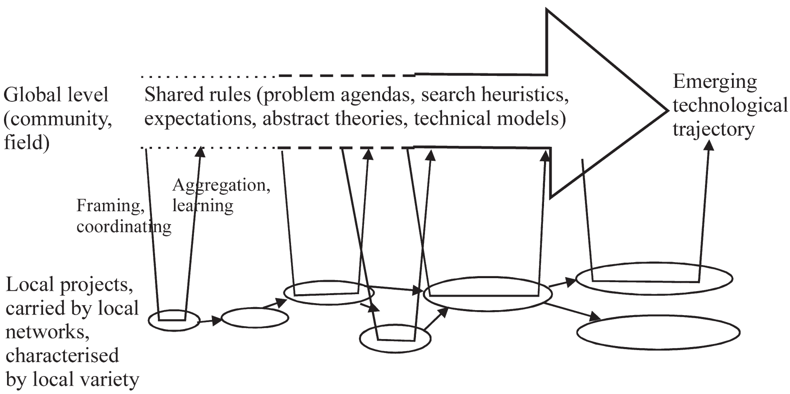
Another MRT is niche theory, which aims to understand the emergence of technical trajectories. Technology life cycle approaches, actor-network theory and social construction of technology have been criticized for portraying the emergence of new technologies as a point-source dynamic, with an empty world being gradually filled up (Russell, 1986; Williams and Edge, 1996). Such point-source dynamics neglect the backdrop of existing technological regimes that are characterized by lock-in, stability and path dependence (Nelson and Winter, 1982; Rip and Kemp, 1998). Niche theory argues that new technologies emerge in small niches outside or at the fringe of existing regimes. Mokyr (1990, 291) characterized new technologies as 'hopeful monstrosities': they are 'hopeful' because product champions believe in a promising future; but they are 'monstrous', because their price/performance characteristics are low. Hence, new technologies cannot survive in mainstream markets and need protection. Niches act as 'incubation rooms', providing space for the nurturing and development of novelties. Niches may have the form of *small market niches*, where selection criteria are different from the existing regime, or they may have the form of *technological niches*, where resources are provided by subsidies.

At the boundary of STS and evolutionary economics, the literature on strategic niche management (SNM) has conceptualized the dynamics of technological niches through an eclectic combination of STS concepts (Schot, 1998; Kemp, Schot, and hoogma 1998). Niche innovations are carried by experimental projects, pilot projects and demonstration projects. In evolutionary terms, such experimental projects provide interaction between variation and selection-environment, allowing users, policy makers and special-interest groups to give feedback to firms, engineers and researchers. Niches thus provide space for interactions between actors and the *building of social networks*. Local projects also provide space for *learning and articulation processes* with regard to technical design, user preferences, regulation, infrastructure requirements, cultural meaning. This allows the alignment of heterogeneous elements in 'configurations that work' (Rip and Kemp, 1998). The third process is the *articulation of expectations and visions*. Expectations attract attention and resources from sponsors, and provide direction to development activities.

While early SNM used these three processes to analyze the success and failure of *individual* projects with new technologies, recent work has looked at sequences of *multiple* projects. This work makes a distinction between local projects and a global niche-level, carried by an emerging field or community (Geels and Raven, 2006). The local-global distinction has social and cognitive dimensions. Law and Callon (1992) distinguished between local and global *networks* in their analysis of an aircraft development project. Local networks refer to actors who are directly involved in projects on location. Global networks consist of actors who have some distance to local projects, but provide resources (finance, political support, technical specifications) that generate a space in which local actors can work. The global network refers to sponsors and a community level. The local-global distinction also has a *cognitive* dimension, distinguishing between specific local knowledge (skills, hands-on-experiences) and abstract, generic, global knowledge shared within a community (Hård, 1994; Deuten, 2003; Geels and Deuten, 2006). The latter are shared cognitive rules (abstract theories, technical models, formulas, guiding principles, and exemplars). Expectations and visions are a special set of cognitive rules that are oriented toward the future and give direction to development activities. Global cognitive rules guide and form resources for activities in local practices but leave room for local interpretations and adjustments. So there can be variety and contingency in local practices, and a degree of stability at the global level, in the form of cognitive rules shared in an emerging community.

The emergence of technological trajectories, which are situated at the global level, takes place through sequences of local projects that gradually add up (Figure 3). Initially, global niche rules and expectations are diffuse, broad and unstable, leaving much room for local improvisation. Local outcomes and experiences are gradually transformed into generic cognitive

Figure 3
Technical Trajectory Carried by Local Projects



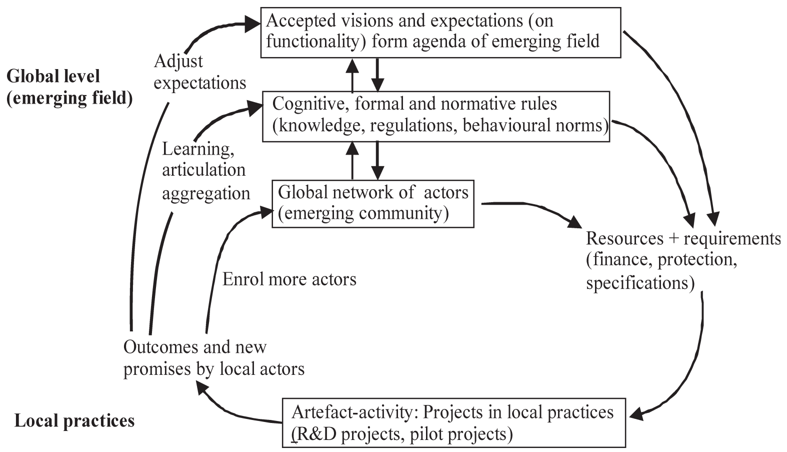
Source: Geels and Raven (2006, 379).

rules through dedicated 'aggregation activities' such as standardization, codification, model building, and formulation of best practices (Deuten, 2003). As global rules become more articulated, specific and stable, the convergence of activities leads to a technological trajectory.

This pattern can be elaborated and explained with an analytical model of interacting niche processes. Figure 4 shows how learning processes, social network building and expectation dynamics influence each other. Actors, embedded in networks, are willing to invest resources in projects, if they have a shared, positive expectation. This shared expectation, together with shared cognitive rules, also provides direction to the projects. Projects, carried by local networks, lead to experiences and outcomes. Local outcomes and learning processes can be aggregated into generic rules. Outcomes are also used to adjust previous expectations and enroll more actors to expand the social network.

This model can also distinguish and explain other patterns in the emergence of new technologies. Garud and Karnøe (2003), for instance, found a breakthrough pattern for wind turbine development in the United States and a bricolage pattern for Denmark. These patterns can be explained by differences in the niche processes. In a breakthrough pattern, learning occurs primarily through R&D and less through local, experimental projects; actors share a search heuristic of rapid technical upscaling; and the

Figure 4
The Underlying Dynamics of Niche Development Trajectories



Source: Geels and Raven (2006, 379).

network is relatively narrow, dominated by engineers, researchers and policy makers. In a bricolage pattern, learning occurs both through R&D and local projects; search heuristics emphasize small, technical steps and solving of bottlenecks; the network is broad, encompassing technology development actors, users, special-interest groups, policy makers. In sum, although SNM is a simple and stylized model, it generates both patterns and explanations.

Multi-Level Perspective on Sociotechnical Transitions

A third MRT example is the multi-level perspective (MLP) on sociotechnical transitions, shifts from one sociotechnical system to another at the level of societal functions. An example for transport systems is the shift from horse-drawn carriages to automobiles. This shift not only involved artifacts, but also infrastructures, regulations, cultural changes, mobility patterns and markets (Geels, 2005a). Another example is the shift in water supply from wells to piped water systems (Geels, 2005b). While many STS scholars focused on the *emergence* of new technologies and systems, less attention has been given to the *shift* from old to new sociotechnical systems. This lack of attention is unfortunate, because it leaves the topic to economists and business scholars, who narrow it down to technological discontinuities, Schumpeterian waves of creative destruction, and the struggle between incumbent firms and new entrants.

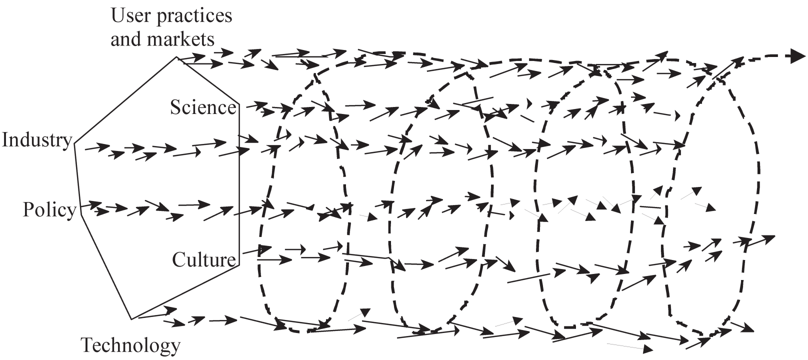
Building on Rip and Kemp (1998), a multi-level perspective (MLP) has been developed to provide a more differentiated sociotechnical understanding of transitions (Geels 2002; 2005a; 2005b, 2005c; Van Driel and Schot, 2005). The MLP follows Misa's (1994, 140-141) suggestion that "a focus on meso-level institutions and organizations that mediate between the individual and the cosmos … offers a framework for integrating the social shaping of technology and the technological shaping of society".

The MLP distinguishes three analytical levels: the niche-level that accounts for the emergence of new innovations, the sociotechnical regime level that accounts for the stability of existing systems, and the sociotechnical landscape that accounts for exogenous macro-developments. The *sociotechnical regime* is an extended version of Nelson and Winter's (1982) technological regime, which refers to cognitive routines shared in a community of engineers. These shared routines guide their R&D activities in similar directions, leading to development along technological trajectories. Sociotechnical regimes are broader, emphasizing that 'relevant social groups' encompass engineers as well as policy makers, users, scientists, media, firms, and special-interest groups (Bijker, 1995). As analytical heuristic, six dimensions are distinguished: technology, science, policy, user practices and markets, cultural meaning, industry and production networks.

Existing sociotechnical regimes are characterized by path dependence, because of several stabilizing mechanisms: favorable regulations and standards, cognitive routines that blind engineers to developments outside their focus (Nelson and Winter, 1982), vested interests and support from incumbent actors, adjustment of lifestyles to technologies, sunk investments in machines, infrastructures and skills. These stabilizing mechanisms do not imply inertia, but incremental innovation and trajectories on the different dimensions of sociotechnical regimes (Figure 5). The different trajectories interact and co-evolve.

This alignment between trajectories creates stability. Fluctuations in one trajectory (political cycles, business cycles, cultural movements, industry lifecycles) are usually dampened by linkages to other trajectories. At times, however, such fluctuations may result in mal-adjustments and tensions, creating windows of opportunity for regime change. Hence, Freeman and Louçă (2001) argue that "It is essential to study both the relatively independent development of each stream of history and their interdependencies, their loss of integration, and their reintegration" (p. 127).[9]

Figure 5
Alignment of Ongoing Processes in a Sociotechnical Regime



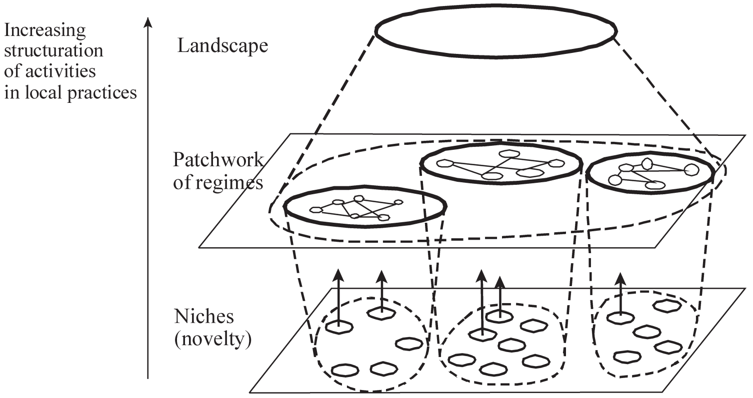
Source: Geels (2005c, 23).

The niche level has been discussed above. The sociotechnical landscape is the third level, which forms an exogenous environment that cannot be influenced directly by regime and niche actors. Examples are macro-political constellations, shared cultural beliefs, natural environment, and demographic developments. The 'landscape' metaphor is used because of the literal connotation of relative 'hardness' and to include the material aspect of society, e.g., the material and spatial arrangements of cities, factories, highways, and electricity infrastructures (Rip and Kemp, 1998).

The (socio)logic of these three levels is that they provide different degrees of structuration to activities in local practices. In niches, structuration is relatively weak, because networks are precarious and cognitive rules are diffuse and unstable. Actors need to put in a lot of 'work' to uphold the niche. In sociotechnical regimes, structuration is stronger. It is possible to deviate from regime-rules, but this is difficult, and takes a lot of effort. Sociotechnical landscapes are hard to deviate from, providing even stronger structuration. The three levels form a nested hierarchy with regard to local practices (Figure 6). Niche-actors *hope* that novelties will eventually be used in the regime or even replace it (hence the arrows in Figure 6). This is not easy, because the existing regime is stabilized and entrenched in many ways.

The multi-level perspective argues that transitions come about through the alignment of processes at different levels. Changes at the landscape level may create pressure on the regime. In response, tensions and de-alignments

Figure 6
Multiple Levels as a Nested Hierarchy



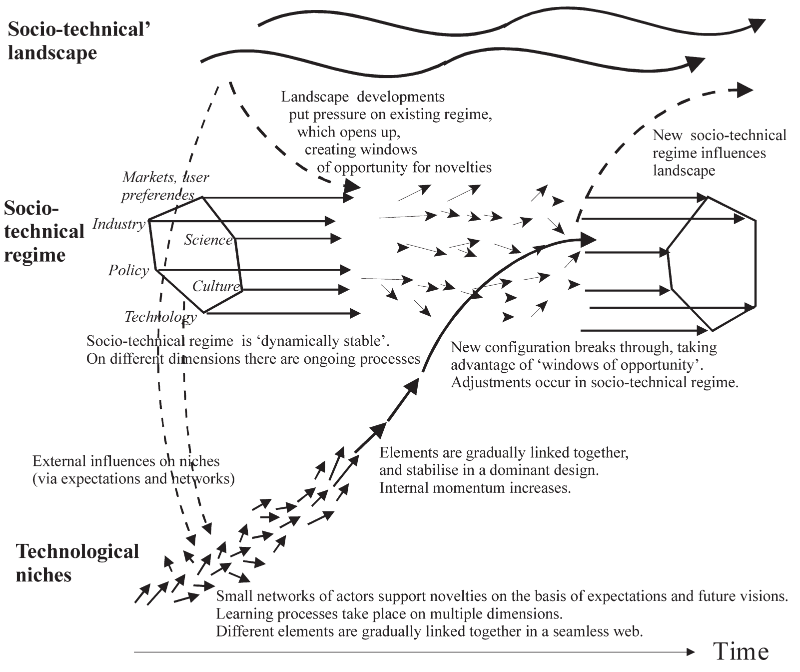
Source: Geels (2002, 1261).

may occur in sociotechnical regimes: e.g., technical problems (bottlenecks, reverse salients), changes in cultural values and public opinion that make side-effects of existing technologies increasingly unacceptable, stricter regulations that cannot be met by the existing technology, changes in markets and user preferences. Such regime tensions create windows of opportunity for wider change. Niche innovations can take advantage of such opportunities when they have stabilized (e.g., a dominant design), improved their price/performance ratio and widened their support network.

If such niche-internal developments link up with regime and landscape processes, novelties may diffuse and compete with the existing technology. This competition is played out in mainstream markets, but also on other sociotechnical dimensions, e.g., lobby for favorable regulations, struggle for supportive research programs, creation of infrastructure, struggle about symbolic meaning. The diffusion of a new technology is thus accompanied by wide changes in the sociotechnical regime (Figure 7).

The MLP analyses existing regimes not (only) as a 'barrier' to be overcome (which would lead to heroic David versus Goliath storylines). Ongoing processes in the regime are important, because they provide opportunities for linking with niche innovations. So the multi-level perspective does away with simple causality and point-source dynamics (as in technology life cycle approaches). There are no single causes of transition, but processes at multiple dimensions and levels simultaneously. Transitions

Figure 7
Multi-level Perspective on Transitions



Source: Adapted from Geels (2002, p. 1263).

come about when these processes link up and reinforce each other. While basic STS notions such as alignment, heterogeneity, and circular causalities can thus be applied, the complexity is reduced to stylized patterns.

The 'linkages between processes at different levels' do not come about automatically, but are made by actors in their cognitions and activities. Admittedly, actors are not very visible in the schematic representations. But agency is implicitly included through the Y-axis ('increasing structuration of activities in local practices'). So actors bring about the different arrows in Figures 6 and 7. Some social groups have much to lose in transitions; others hope to gain. Because actors have different perceptions and interests, they struggle, negotiate, and form coalitions during transitions. Hence, transitions are contested and emergent outcomes of interactions between social groups with myopic views and differing interests, strategies and resources.

These dynamics are important and can be analyzed in detailed case studies. Despite this empirical complexity, which differs for each case study, the multi-level perspective shows that aggregate patterns can be identified.

Concluding Remarks

How do these examples relate to the problems sketched in the introduction? Regarding the first problem (policy relevance), promise-requirement cycles inform new large-scale research programs in nanotechnology and genetics. These programs have allocated percentages of their budget to technology assessment projects, not in the (old) sense of forecasting impacts, but to enhance the quality and robustness of promises that inform technical activities. Strategic niche management and the multi-level perspective also have policy relevance, especially in the context of Dutch environmental policies. The Ministry of Energy uses SNM as a frame to better manage portfolios of pilot projects with renewable energy technologies (focus more on sequences of projects and circulation of experiences). The Dutch government is also interested in transitions to sustainability, and has adopted the MLP as a frame to think about such long-term regime shifts. Regarding the second problem (unrelated concepts), the three 'cases' are examples of MRT in the sense of explicitly relating concepts into stylized analytical models. Regarding the third problem (too much focus on complexity, local practices, contingency) the examples show that one can abstract from complexities and derive generic patterns. Of course, each concept in the examples can be unpacked and shown to be more complex. Aggregation, for instance, is not simply a cognitive process of drawing general lessons from local projects; one can also show how power, negotiation, and framing play a role. But unpacking and demonstration of complexity are not the only interesting scientific activities. The examples of MRT sacrificed some empirical detail and local complexity for gains in generality and abstracted patterns. In this respect Poole and Van de Ven (1989) made an interesting distinction, arguing that process theories should have two complementing components: global and local models.

> "The global (macro, long-run) model depicts the overall course of development of an innovation and its influences, while the local (micro, short-run) model depicts the immediate action processes that create short-run developmental patterns. . . . A global model takes as its unit of analysis the overall trajectories, paths, phases, or stages in the development of an innovation,

> whereas a local model focuses on the micro ideas, decisions, actions or events of particular developmental episodes" (p. 643).

STS has given much attention to local models. But the examples show that STS can also make interesting and relevant contributions to global models. MRT is an invitation to explore this avenue further. This also addresses the fourth problem (academic styles). While enlightenment, deconstruction, reflexivity, and local complexities remain important in STS, MRT suggests an interesting additional academic style. The examples show that STS can derive patterns and mechanisms, working on description *and* explanation. The promise-requirement cycle is a mechanism that helps explain the pattern of hype-disappointment cycles. Learning processes, network building and expectation dynamics are mechanisms in SNM that help explain the emergence of technological trajectories and particular innovation journeys (breakthrough or bricolage). Only the multi-level perspective is not yet a 'real' MRT. It conceptualizes a pattern of long-term change, but says less about mechanisms. At present, the MLP is only a global model. More work should be done on mechanisms. We are presently working on distinguishing different transition *pathways*, depending on differences in timing and kinds of multi-level interactions and on different interactions between social groups (Geels and Schot, 2007).

A further reflection is that the examples address longer time-frames (decades) and larger scales (communities, sociotechnical systems) than much STS work, which often focuses on shorter-term technical projects and local practices. The article thus responds to Pestre's (2004, 360) challenge: "My feeling is that we have arrived at a moment where we should rethink SSK's asceticism, radical agnosticism and de facto scientism, and return, at least some of the way, to more macro approaches and questions. In order to do this, we will have to abandon the idea of privileging the mere *following of actors* and introduce more explicitly *various scales* of analysis" (italics in original). MRT forms an interesting response to this call. STS is well-placed to pursue this avenue, because there are already achievements in that direction.[10]

Acknowledgements

I thank Sally Wyatt, Brian Balmer and four anonymous referees for their useful comments on previous versions of this article. I also gratefully acknowledge financial support from the Dutch Knowledge Network on System Innovation (KSI).

Notes

1. This article is part of a special issue on middle-range theory and STS edited by Brian Balmer and Sally Wyatt

2. One grand theory school was classical organization theory (scientific management, bureaucratic theory, management principles), another was the human relations movement.

3. For some exceptions with a positivistic slant see MRT in nursing theory and archeaology.

4. This is also a matter of style. STS often wants to 'educate' and 'enlighten' policy makers. This desire to share our interesting ideas has somewhat of a supply-push character. Enhanced policy relevance may also depend on taking the user (i.e., policy maker) seriously. Of course, this is not always easy, especially when the user does not know precisely what (s)he wants.

5. "What I am trying to do is to attack simplicity—and a notion of theory that says that it is or should necessarily be simple, clear, transparent. … I believe that they render complex thinking (thinking that is not strategically ordered, tellable in a simple way, thinking that is lumpy or heterogeneous) difficult or impossible" (Law, 1999: 8-9).

6. All "essentialist divisions are thrown on the bonfire of the dualisms: thruth and falsehood, large and small, agency and structure, human and non-human, before and after, knowledge and power, context and content, materiality and sociality, activity and passivity" (Law, 1999: 3).

7. "The ridiculous poverty of the ANT vocabulary … was a clear signal that none of these words could replace the rich vocabulary of the actor's practice. … The weakness on our part does not mean, however, that our vocabulary was too poor, but that, on the contrary, it was not poor enough" (Latour, 1999: 20).

8. "For by now we know that these stories do not necessarily add up, do not necessarily come to a point. That we may give up single narratives in favour of many small stories. … *Perhaps there is no pattern*, no overall pattern. Perhaps there is nothing except practices. Perhaps there is nothing other than stories, performing themselves and seeking to make connections, practical and local connections, specific links" (Law, 1996: 9) (italics in original).

9. To understand long wave dynamics, they look at interactions between dynamics in five sub-systems: science, technology, economy, politics and culture.

10. The examples discussed in this article are not the only achievements. Other patterns and mechanisms have been identified, e.g., the scientific credibility cycle, the sailing ship effect, increasing returns to adoption.

References

Bijker, W.E. 1993. Do not despair: There is life after constructivism. *Science, Technology, & Human Values* 18(1): 113-138.

Bijker, W.E. 1995. *Of bicycles, bakelites and bulbs: Towards a theory of sociotechnical change*. Cambridge, MA: The MIT Press.

Boudon, R. 1991. What middle-range theories are. *Contemporary Sociology* 20(4): 519-522.

Brown, N., and M. Michael. 2003. The sociology of expectations: Retrospecting prospects and prospecting retrospects. *Technology Analysis & Strategic Management* 15(1): 3-18.

Callon, M. 1986. The sociology of an actor-network: The case of the electric vehicle. In *Mapping the dynamics of science and technology*, edited by M. Callon, J. Law and A. Rip, 19–34. London: MacMillan

Callon, M., and B. Latour. 1992. Don't throw the baby out with the Bath school: A reply to Collins and Yearley. In *Science as practice and culture*, edited by A. Pickering, 343-368. Chicago and London: The University of Chicago Press.

Collins, H.M., and S. Yearley. 1992. Journey into space. In *Science as practice and culture*, edited by A. Pickering, 369-389. Chicago and London: The University of Chicago Press.

Collins, H.M., and R. Evans. 2002. The third wave of science studies: Studies of expertise and experience. *Social Studies of Science* 32(2): 235-296.

Davis, J.A. 1994. What's wrong with sociology? *Sociological Forum* 9(2): 179-197.

Deuten, J.J. 2003. *Cosmopolitanising technology: A study of four emerging technological regimes*. Enschede: Twente University Press.

DiMaggio, P.J. 1995. Comments on "What theory is not". *Administrative Science Quarterly* 40(3): 391-397.

Edge, D. 2003. Celebration and strategy: The 4S after 25 years, and STS after 9-11. *Social Studies of Science* 33(2): 161-169.

Flen, J., and A. Linden. 2005. Gartner's hype cycle special report for 2005. August 5, 2005. Gartner Research. Available from http://www.gartner.com/resources/130100/130115/gartners_hype_c.pdf (accessed June 1, 2007).

Freeman, C., and F. Louçă. 2001. *As time goes by: From the industrial revolutions to the information revolution*. Oxford: Oxford University Press.

Garud, R., and P. Karnøe. 2003. Bricolage versus breakthrough: Distributed and embedded agency in technological entrepreneurship. *Research Policy* 32(2): 277-300.

Geels, F.W. 2002. Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. *Research Policy* 31(8/9): 1257-1274.

Geels, F.W. 2005a. The dynamics of transitions in socio-technical systems: A multi-level analysis of the transition pathway from horse-drawn carriages to automobiles (1860–1930). *Technology Analysis & Strategic Management* 17(4): 445-476.

Geels, F.W. 2005b. Co-evolution of technology and society: The transition in water supply and personal hygiene in the Netherlands (1850–1930)—a case study in multi-level perspective. *Technology in Society* 27(3): 363-397.

Geels, F.W. 2005c. *Technological transitions and system innovations: A co-evolutionary and socio-technical analysis*. Cheltenham: Edward Elgar.

Geels, F.W., and W.A. Smit. 2000. Failed technology futures: Pitfalls and lessons from a historical survey. *Futures* 32(9/10): 867-885.

Geels, F.W., and R.P.J.M. Raven. 2006. Non-linearity and expectations in niche-development trajectories: Ups and downs in Dutch biogas development (1973-2003). *Technology Analysis & Strategic Management* 18(3/4): 375-392.

Geels, F.W., and J.J. Deuten. 2006. Local and global dynamics in technological development: A socio-cognitive perspective on knowledge flows and lessons from reinforced concrete. *Science and Public Policy* 33(4): 265-275.

Geels, F.W., and J.W. Schot. 2007. Typology of sociotechnical transition pathways: Refinements and elaborations of the multi-level perspective. *Research Policy* 36(3): 399-417.

Guggenheim, M., and H. Nowotny. 2003. Joy in repetition makes the future disappear: A critical assessment of the present state in STS. In *Social studies of science & technology: Looking back, ahead*, edited by B. Joerges and H. Nowotny, 229-260. Dordrecht: Kluwer.

Hård, M. 1993. Beyond harmony and consensus: A social conflict approach to technology. *Science, Technology, & Human Values* 18(4): 408-432.

Hård, M. 1994. Technology as practice: Local and global closure processes in diesel-engine design. *Social Studies of Science* 24(3): 549-585.

Hedström, P., and R. Swedberg, eds. 1998. *Social mechanisms: An analytical approach to social theory*. Cambridge: Cambridge University Press.

Kemp, R., J. Schot, and R. Hoogma. 1998. Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management. *Technology Analysis & Strategic Management* 10(2): 175-196.

Latour, B. 1999. On recalling ANT. In *Actor network theory and after*, edited by J. Law and J. Hassard, 15-25. Oxford, UK: Blackwell.

Law, J. 1992. Notes on the theory of the actor network: Ordering, strategy and heterogeneity. *Systems Practice* 5(4): 379-393.

Law, J., and M. Callon. 1992. The life and death of an aircraft: A network analysis of technical change. In *Shaping technology/building society: Studies in sociotechnical change*, edited by W.E. Bijker and J. Law, 21-52. Cambridge, MA: The MIT Press.

Law, J. 1996. Traduction/trahison: Notes on ANT. *Online paper from Actor-Network Theory Resource* (Available from http://www.lancs.ac.uk/fass/centres/css/ant/index.htm) (accessed January 8, 2007).

Law, J. 1999. After ANT: Complexity, naming and topology. In *Actor network theory and after*, edited by J. Law and J. Hassard, 1-14. Oxford, UK: Blackwell.

Merton, R.K. 1948. Discussion of Parsons' 'The position of sociological theory'. *American Sociological Review* 13(2): 164-168.

Merton, R.K. 1968. *Social theory and social structure*. Third edition (previous editions in 1949 and 1957), Glencoe, IL: Free Press.

Mills, C.W. 1959. *The sociological imagination*. London, Oxford, New York: Oxford University Press.

Misa, T.J. 1994. Retrieving sociotechnical change from technological determinism. In *Does technology drive history? The dilemma of technological determinism*, edited by M.R. Smith and L. Marx, 115–141. Cambridge, MA: The MIT Press.

Mokyr, J. 1990. *The lever of riches*: *Technological creativity and economic progress*. New York: Oxford University Press.

Molina, A. 1995. Sociotechnical constituencies as processes of alignment: The rise of a large-scale European information technology initiative. *Technology in Society* 17(4): 385-412.

Nelson, R.R., and S.G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Belknap Press.

Perez, C. 2002. *Technological revolutions and financial capital: The dynamics of bubbles and golden ages*. Cheltenham: Edward Elgar.

Pestre, D. 2004. Thirty years of science studies: Knowledge, society and the political. *History and Technology* 20(4): 351-369.

Pinder, C.C., and L.F. Moore, eds. 1980. *Middle range theory and the study of organizations*. Boston: Martinus Nijhoff.

Poole, M.S., and A.H. Van de Ven. 1989. Towards a general theory of innovation processes. In *Research on the management of innovation: The Minnesota studies*, edited by A.H. Van de Ven, H.L. Angle, and M.S. Poole, 637-662. New York: Harper & Row Publishers.

Rip, A., and R. Kemp. 1998. Technological change. In *Human choice and climate change*, Volume 2, edited by S. Rayner and E.L. Malone, 327-399. Columbus, OH: Battelle Press.

Russell, S. 1986. The social construction of artifacts: A response to Pinch and Bijker. *Social Studies of Science* 16(2): 331-346.

Schot, J.W. 1998. The usefulness of evolutionary models for explaining innovation. The case of the Netherlands in the nineteenth century. *History of Technology* 14: 173-200.

Van Driel, H., and J. Schot. 2005. Radical innovation as a multi-level process: Introducing floating grain elevators in the port of Rotterdam. *Technology and Culture* 46(1): 51-76.

Van Lente, H., and A. Rip. 1998. Expectations in technological developments: An example of prospective structures to be filled in by agency. In: *Getting new technologies together*, edited by C. Disco and B.J.R. van der Meulen, 195-220. Berlin and New York: Walter de Gruyter.

Weick, K.E. 1999. Theory construction as disciplined reflexivity: Tradeoff in the 90s. *Academy of Management Review* 24(4): 797-806.

Williams, R., and D. Edge. 1996. The social shaping of technology. *Research Policy* 25(6): 865-899.

Woolgar, S. 2004. What happened to provocation in science and technology studies? *History and Technology* 20(4): 339-349.

Frank W. Geels is assistant professor at Eindhoven University of Technology, the Netherlands. His interests focus on radical innovation, large-scale technological transitions and system innovations. Insights from sociology of technology, innovation studies, history of technology and evolutionary economics are used to understand the underlying dynamics. The conceptual work is grounded with case studies.