FOUNDATIONS AND FUTURES OF INNOVATION MANAGEMENT THEORY

Rasmus Koss Hartmann

Foundations and Futures of Innovation Management Theory

Foundations and futures of innovation management theory

Version 0.11. Updated September 15, 2018.

Rasmus Koss Hartmann

Center for Integrative Innovation Management

University of Southern Denmark

Contents

Preface

I always disliked textbooks, both as a student and as a teacher. In that sense, it is more than a little paradoxical to be writing one. What I always disliked was the sense they gave of being overly comprehensive and somehow final, like this was the 'end of history' of a topic. Their main purpose, it seemed to me, was often to present a seemingly unending list of more or less stylized facts, reducing the complex arguments of a research paper to a single sentence in the process.

I would later be told that 'research is a conversation', which made me think back to textbooks as particularly bad at capturing that sense of both disagreement and dynamism. Later, I would also read about how technology either substitutes or complements labor and, with some categorical slippage, again think back to textbooks as exemplary of (often poor) substitutes for reading of source ideas. And yet, for all these problems, students would often ask for textbooks because they understandably found it to be a bit of dive into the intellectual deep end to grapple directly with research papers. This is understandable. Research papers are written by academics for academics, and so they take things for granted about their reader that it is blatantly unfair to take for granted about students even at advanced levels, and even more unfair for students just beginning to engage with a topic area. Of course, this does not imply that we should consider papers somehow beyond the reach of students. I am a big proponent of challenging students *much* more than business school classes typically do, but also believe that such challenge should be accompanied by lots of support.

And thus, the question that emerged was how to go about putting such a book together, how to build a textbook of sorts that complemented rather than substituted direct engagement with challenging material. What you have in front of you is the first answer to that question. More precisely, it is a beta-version of an answer and one that will continually evolve as students help me think more and better about progressively better answers to the question.

The primary purpose of this book is to provide a systematic introduction to innovation management theory, but to do so in a way that brings readers *closer* to the source ideas. It is intended to support students in readings of original texts. It is a hybrid between the conventional textbook and a curated volume of selected readings. It is, in a sense, an elaborate version of what could be called a 'reader' in innovation management theory. It is meant to be used alongside those original texts to make them more accessible and to sort the textual and ideational wheat from the chafe. Despite academic authors' best efforts, there is often a lot of chafe in academic writing (my own certainly included) owing to genre conventions, jargon, signposting and many other things.

The secondary purpose is to convey a particular image of innovation theory as a body of knowledge with many interconnections. Some of those are explicit and when you read you will see references back and forth between the ideas in the various chapters. Other interconnections can be created through theoretical analysis. We can, so to speak, use texts to make each other better and more interesting. While interconnected, ideas are not all in agreement and I do not intend (as textbooks often do) to create a sense of finality. Hopefully, what will emerge will be a sense of where in the literature there are connections, continuities, disjuncture, agreements and unfinished discussions. Rather than include the latest and greatest empirical results and their attendant weighing-in on discussions, I have chosen to emphasize theoretical ideas that prepare students to read and make sense of empirical discussions in more advanced classes.

It is worth stating the somewhat obvious at the outset. As has to be the case, this book contains both a select set of ideas and a select reading of those ideas. They are my selections and reflect both my understanding of innovation research as a whole and my theoretical priors. In selecting between possible ideas to include, I have *largely* tried to choose ideas that represent a consensus view of what are canonical concepts, the foundational pillars of innovation management theory that most professional innovation scholars would consider ‘mainstream’ and, we would like to believe, relatively enduring. To that end, we will be touching quite a lot of disciplinary ground. We will be moving from the *economics of innovation* through *economic history* over *science and technology studies*, *organization and management studies* (of both the sociological and economic variety), *strategy* and *psychology*, and will have occasional brushes with geography on the way. That does not make for a particularly straightforward story, and that is fine.

In presenting my reading of those ideas, I have on occasion deliberately tried to be less consensus-oriented and less mainstream. Essentially, my point of view is informed by two contrarian theoretical traditions, namely those of Critical Management Studies and Austrian Economics. Both make a virtue of criticism and disbelief in many of the central tenets of mainstream theory and mainstream empirical methods.

That disbelief will shine through, sometimes brightly, in the chapters that follow. This means that my readings tend to be more skeptical of statistical analyses, of the potency of strategies and managers to analyze and to affect change and of the prescriptive power of theory than you might generally find in innovation research. Relative to mainstream researchers, I tend to be more (but not necessarily) persuaded by historical studies and up-close empirical observation, more concerned about things like power, ideology and knowledge and less optimistic about the ability of management theory to provide ‘science-y’ answers for management practice. I have included a presentation of both perspectives in Appendix I, such that readers who (completely reasonably) are not familiar with the ideas can read my readings in an informed light.

This balance between the mainstream and the critical provides, I hope, a somewhat more interesting and more conversational access point to the ideas of the book, because they get us away from the common problem in textbooks of presenting things as indisputable. While I will continually try to connect the foundational ideas to each other, I will also try to problematize them and dispute them, to use them as sounding boards for reflection or as a repertoire of potential interpretations that might or might not be relevant in understanding particular contexts – theories’ implications for practice, I believe, depends on the practice. While some of the texts you will read are interested in the predictors of ‘performance’, you will not really see me present them as ‘recipes’ for that performance. And I will be as explicit as I can about when I am stating generally held views and my own, less generally held views.

Also to be stated at the outset is my intention for the book’s use. There are several ways to use the book, depending on ambition and how one makes use of the podcasted lectures that are meant to supplement the book. At the most basic level, you can listen to the podcasts as a (somewhat heavy) alternative to Spotify on your commute, as ‘edutainment’. You can also read this book together with the podcasts, in which case you have something resembling the conventional combination of a textbook and lectures that is typical of bachelor’s level university courses, but without the requirement of buying a textbook and attending a set of classes. You could also just read this book in isolation. That will surely at a minimal level give you an impression of innovation management as a topic – an impression that can only be better than those hit-and-run impressions that you will get from the shelves of airport bookstores that tend to spill over with innovation management advice.

The way the book is designed is for you to use the book, possibly together with its attendant podcasts, as an entry point to engaging with some (or all) of the texts that the individual chapters cover. When I teach using this book, I do it according to the model described in appendix II which (unsurprisingly perhaps) I think is the most demanding and *by far the most rewarding* way to use it. That is how I encourage it to be used, but I am always happy to see when new ideas and applications grow out old ones. That, after all, is the curious and wonderful way that innovation over time makes us all better off.

Chapter I:
The challenge of innovation under capitalism

Texts in this chapter:

Schumpeter, J. A. 1942 (2008). **Capitalism, Socialism and Democracy** (Chapter 8). Harper Perennial Modern Thought Edition.

McCloskey, D. N. 2010. **Bourgeois Dignity: Why Economics can't explain the modern world** (Chapter 1). University of Chicago Press.

Abernathy, W. J. 1978. **The productivity dilemma: Roadblock to innovation in the automobile industry** (chapter 4)**.** John Hopkins University Press.

March, J. G. 1991. **Exploration and exploitation in organizational learning.** Organization Science. 2(1), 71-87.

In this chapter, we are going to set up the basis for the remainder of the book, the most foundational of foundational understandings. This basis is going to be the dual challenge that firms face with regards to innovation. This is the challenge: firms need to compete *here and now* but they also need to compete *long term* in order to not get pushed aside by new firms with new technologies. We are going to look into the phenomenon that firms *do* get pushed aside and I will be introducing a widely accepted argument that the reason they get pushed aside is because they get 'out-innovated'. From this come an imperative for firms to innovate in order to succeed in the long term. This innovation imperative is, also uncontroversially, widely accepted as the key driver of increased wealth and living standards globally and especially in the industrialized West.

The innovation imperative, however, creates a special kind of challenge for firms – what is widely known as the productivity dilemma. This dilemma in its essence is that the conditions that enable firms to innovate (which is necessary to compete in the long term) are very different from those that enable firms to be highly efficient (which is necessary to compete in the short term). Dealing with that dilemma is the foundational problem – both practically and theoretically – of innovation management, perhaps even of business strategy, and that is the dilemma to which we will address ourselves throughout the book.

1.1. Schumpeter: Creative Destruction

The signs of Creative Destruction are all around us. As consumers, we are incessantly presented with new products, services and solutions. Things that were once stables of everyday life disappear. We regularly find out that we need things that were until recently nonexistent, even unimaginable. What we recognize as the largest companies, those that most dramatically shape our consuming lives by providing us with all this novelty, are mostly new ones. Just a generation ago, many of them either had not been founded or were in their absolute infancy. The technologies that these companies use similarly did not exist a generation ago. The same can be said of the technologies that shape our working lives. More and more of us work or will work in jobs that a generation ago made up a small share of the overall workforce or simply were not invented. We will also fulfill those jobs under organizational arrangements that are new. If you live in a Western city, you will see what were once factories being turned into fashionable housing and offices for a new kind of work and for a new kind of production. This churn, this constant change, is Creative Destruction unfolding.

Schumpeter's book *Capitalism, Socialism and Democracy* was written in 1942, amidst the horrors of a World War and at a time when many were deeply disillusioned with capitalism as an economic system. While many today share the disillusionment with capitalism, the real difference between then and now is that in the 1930s and 1940s socialism really was an alternative, with a substantial share of the world's population living in socialist economies. At the time, many looked to the the Soviet Union and saw a well-functioning economic model, while in the West they would see something not quite undesirable. They would often *not* see the free markets and perfect competition on which capitalism is predicated, but monopolies and think (with Adam Smith in mind) that this had to be economically inefficient and (with Karl Marx) that the tendency towards monopolies was inherent in capitalism, as was exploitation of the working class. If you looked, you might see the same today. They would also see vast inequality, with incredible wealth held by what we would call the 1% and the 99% have just suffered through a desperate Depression.

Schumpeter's book was an effort to understand socialism and capitalism and their respective relationships with democracy. In what is perhaps the most concise characterization of capitalism ever written, this would be Schumpeter's argument: "Creative Destruction is the essential fact about capitalism. It is what capitalism consists in and what every capitalist [firm] has got to live in."

To understand that, we need to understand how Schumpeter thought about capitalism. "The essential point to grasp", he writes, "is that in dealing with capitalism we are dealing with an evolutionary process… Capitalism, then, is by nature a form or method of economic change and not only never is but never can be stationary" (p. 82). Capitalism is inherent unstable and dynamic. To live in a capitalist economy is to live in circurmstances that, for better and worse, constantly change.

There is a deep poetry, I think, to Schumpeter's account of this constant change. Creative Destruction is a very evocative and quite intuitive way to capture something that is truly profound: that capitalism implies a constant 'industrial mutation' that "incessantly revolutionize the economic structure *from within* [italics in original], destroying the old one, incessantly creating a new one". To him, there is a storm, a 'perennial gale of Creative Destruction' blowing through capitalist society, like the Hindu god Shiva dancing through the world, tearing down what is and letting the new emerge from the ashes.

While we may not appreciate the causes of this process or the profundity of its implications, the *threat* of Creative Destruction is keenly felt by businesses. Schumpeter talks about Creative Destruction as an 'ever-present threat' (p. 85), one that 'disciplines before [new competition] attacks' (p. 85). It creates a paranoid anxiety for firms that they must compete not just in the short term on price and quality by constantly marginally improving. This is the kind of competition that firms are always in. In the slightly longer term, they also have to deal with "competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits and the outputs of the existing firms but at their foundations and their very lives. This kind of competition is as much more effective than the other as a bombardment is in comparison with forcing a door" (p. 84-85). The threat of such Creative Destruction constantly looms, even over monopolies, and makes them act *as if* they were in a state of cutthroat competition, because even if they are not currently competing they still compete against potential future entrants to their industry that will seek to overtake their position.

For Schumpeter, *innovation* is the fundamental impulse spurring Creative Destruction, the force that propels change in the economy. He employs a relatively broad definition of innovation, covering "new consumer' goods, new methods of production or transportation, the new markets, the new forms of industrial organization" (p. 83). Today, we tend to use that as the definitional starting point for distinguishing between different forms of innovation, rather than use the umbrella term, but do bear in mind that innovation can take other forms than those presented here. *Product innovation* is what people typically think about when they think of innovation, because it refers to either new physical products or (increasingly) services. These tend to be highly visible. *Process innovation* covers new techniques or tools for creating products or services. These tend to be less visible and attention-grapping, because they often are sold in business-to-business markets, but are indisputably important. *Organizational innovation* covers new ways to organize work and production. It seems like a natural thing today, but the corporation as we know it today did not always exist, but was actively invented. It is an innovation, as are many other facts of organizational life. *Market creation* refers to the opening up of new markets, whether by extending the geography in which a solution is marketed or in extending technologies into new fields. Innovations in transportation get relatively little attention in most modern innovation research, but is nonetheless profound in its consequences. Consider, for instance, the implications that the humble 40-foot container has had on global trade and as a consequence on what you consume (Lewinson, 2008). Some of the very meaningful additions that we could make to the list are innovation in business models (e.g. new ways create and capture value) or techniques (e.g. new ways to *use* products).

These sub-classifications are worth noting in part for their own sake – as we will return to, they each have their own set of causes and a very different set of effects on organizations. Contrary to what many would have you believe, being good at one does not imply being good at the other, as we will return to in the essays that follow.

The other important thing to understand about these five forms of innovation is that they in interaction create what Schumpeter would surely see (and what you should see) as the miracle of capitalism. They allow for the long-run expansion of output and the driving down of prices, putting goods and services that were once only accessible to the hyper-rich into the hands of all of us. Innovation is what allows those of us who are fortunate enough to be born into the affluent societies of the Western world to enjoy standards of living that were frankly unimaginable just 200 years ago. To be clear, capitalism has incredible and far-reaching downsides. It creates gaping holes in our souls as both workers and consumers (as we will return to) and is a driving force in creating ecological damage on a potentially catastrophic scale. But be clear also that the kind of growth in material wealth that capitalism has created is a miracle that is unprecedented in world history, as Deirdre McCloskey's work will make eminently clear in the next essay.

Going forward, we are going to see Creative Destruction come up again and again as a mainstay of innovation management theory. We will see it as the constant threat that Schumpeter describes. In this capacity, all the various ways that established firms and industries could be destroyed by innovation weigh heavily (and should weigh heavily) on the minds of incumbent firms. We will see the dilemma that it poses between competing at the margins *and* at the center of profits as constitutive of one of the key challenges of innovation management: of balancing exploration and exploitation both at the level of the firm allocating its resources between different activities and at the level of how a technology should be developed. We will also see it as an opportunity. Just as innovation can destroy, it can also elevate new entrants and new ideas to prominence. Attackers can have advantages over incumbents and Creative Destruction is something that people can actively try to bring about. We will see the sense of turmoil that Schumpeter describes as a stable of industrial change, but also as something that firms try to bring to an end. We will see the tension between the immaculate wealth creation and the far-reaching and painful destruction as a key challenge facing society. And we will see how the constant flux, the constant search for new and better alternatives is currently re-shaping the way innovation happens and firms organize. We live, for better and for worse, in fantastically interesting times and understanding innovation and its implications is a central part of understanding them.

1.2. McCloskey: The origins of capitalism as we know it

Contrary to what many experience today, human history is – in material terms – mostly a story of *non-change*. Then, germinally around the end of the 18th century and forcefully through the Industrial Revolution and up to today, everything changed. McCloskey makes this evocatively clear: "Economic history has looked like an ice-hockey stick lying on the ground. It had a long, long horizontal handle at $3 a day extending through the two-hundred-thousand-year history of *Homo Sapiens* [italics in original] to 1800, with little bumps upward on the handle in ancient Rome and the early medieval Arab world and high medieval Europe, with regression to $3 afterward – then a wholly unexpected blade, leaping up in the last two of the two thousand centuries, to $30 a day and in many places well beyond" (p. 2). We are fortunate to live on that blade, amid the miracle of capitalism. As McCloskey makes clear, for all the very real ailments of modern capitalism, the world we live in is undisputedly *much better* for *a greater share* of people than has been the case at any other point in human history. This is the case because of *economic growth*.

In his book *The Lever of* Riches, economic historian Joel Mokyr argues that economic growth in a society can come from several sources. We can grow richer through 'Smithian growth' (named for Adam Smith, who needs no introduction), which comes from increased trade. Increased trade allows for larger markets and for increased specialization and with that comes increased productivity, which makes society richer. We can also have 'Solovian growth' (named for Robert Solow, an American economist working in the second half of the 20th century). This comes from increased investment in what economist call capital goods, i.e. production equipment. Having more production equipment allows workers to be more productive and society to produce more goods with the same inputs. To this McCloskey might add 'Marxian growth' (after Karl Marx), which would come from increasingly brutal exploitation of workers. Just as having more production equipment can make workers more efficient, pushing them harder can increase productivity, but only up to a certain back-breaking point. These things surely matter in explaining why we are richer today than we were in the past, but they *cannot* as McCloskey argues explain the blade of the hockey stick. That change is too momentous to be explained by trade, capital accumulation, exploitation of workers or any other single concept in the economist's theoretical toolbox. Except, that is, for *one* kind of growth, the kind that Mokyr calls *Schumpeterian growth*: growth in human knowledge and ideas and innovation. Innovation *more than anything else* explains the material wealth of the modern world[1]. That is why we need to understand it. Not just because it is essential for firms to compete effectively, but because innovation is what makes *society* better off.

---

[1] The reason that innovation can explain growth in ways that more material factors cannot relates to the particular economics associated with *ideas*. The short story is that ideas allow for what Mokyr calls 'free lunches', because they can for instance be infinitely replicated at little or no cost, at least in principle. The longer story about the unique economics of ideas, which in many way resemble the economics of digital goods, is well told in Shapiro & Varian's book 'Information Rules' (1998)

The question for McCloskey, then, is what brought about all this innovation? Why did the Industrial Revolution happen and why does it continue to happen? Why did we get on the hockey stick's blade when we did not, say, two hundred years before (in which case we would be unfathomably well-off today) or two hundred years later (in which case you probably would not be reading this book, but toiling away in a field somewhere).

Her answer is that things began to change because *ideology* had been changing[2]. Following the Reformation, new ideas began to emerge about the urban middle class of hiring, owning, professional or educated persons (i.e. the 'bourgeoisie'), first in the Low Countries and later in Britain and ultimately in the United States. To be clear, it was *not* the Reformation and the attendant emergence of what German sociologist Max Weber (1930) calls the 'Protestant work ethic' that created the modern 'the Spirit of Capitalism' (because, if that were the case, why do we see capitalism today in distinctly non-Protestant countries?). It was that with the reformation came a new way of thinking and talking about the bourgeois class. This thinking, in an evolved version, is still with us today and profoundly shapes how people think and talk about innovation in society and organizations.

Where before this group of merchants had been viewed with some contempt, they began to be viewed more favorably, as dignified and not repugnant. Up to this point, the best and brightest had gone to work in the courts, the church or the military, because those were 'honorable' professions. Alas, they are also mostly conservative ones. But with changing attitudes towards merchants, "[i]t became honorable… to invent a machine for making screws or to venture in trade to Cathay" (p. 12). "Ordinary conversations about innovation and markets became more approving" (p, 7), "general opinion shifted *in favor* [italics added] of the bourgeoisie and especially in favor of its marketing and innovation" (p. 7). This process had been going on for a while (that is the way it is with the ebb and flow of the tides of history) before the Industrial Revolution took off. What at the end of the 18th century emerged from that process was that ideology changed in such a way as to make Creative Destruction the socially desirable thing *that it had not been before*. The result of that change in ideology has been a "novel and immense and sustained, almost lunatic, regime of innovation" (p. 19). With this regime arose what Schumpeter in his day and we in ours recognize as capitalism, that dynamic economic system of which Creative Destruction is the essential fact.

To be clear, what was new was neither capitalism nor innovation. Capitalism existed long before the early 19th century, but this was capitalism in the sense of commercial activity, with trade, with prices emergently determined by supply and demand, with financial credit, with money and accumulation of wealth. None of those things, as McCloskey makes totally clear in her subsequent chapters, were new. Innovation also was not new. In Mokyr's book, he describes Western technological progress from Classic Antiquity through to the Middle Ages and the Renaissance before arriving at the Industrial Revolution's 'Years of Miracles'. You could similarly describe an important history of Innovation in ancient China, in the empires of South America up to the brutality of European conquest, in the Arab world and at some level in all human society, ever. Innovation has been with us since our ancestors on the Savannah. McCloskey notes technological change already in the Upper Paleolithic and amongst the proto-Australians (innovations in boats is what got them to Australia, after all), to give just a few of her examples. What was new was the *scale* of innovation and with it a new competitive process at the heart of capitalism. Sometimes, a change in degree is a change in kind and that is exactly the

[2] Do note that there are other answers to this question, such as those advanced by e.g. Rosenberg & Birdzell (1987) or McGraw (1995).

case here. A change in ideology spurred the gale of Creative Destruction to the historically unprecedented heights that we see today.

This image of innovation as ideologically driven is something that I encourage you to take into your further reading. It is also something that we should have some concerns about. There can be little doubt that we are living in a time obsessed with innovation. Just as you can see the signs of Creative Destruction all around you, you will also hear what some call the gospel of innovation preached and hyped by practitioners and scholars in business and management and technology, sometimes with good intentions and other times for personal profit in some form. Change, they say, is the only constant and innovation the only answer. The ideological pendulum has swung very, very far from when change was a dirty word and innovation not yet coined.

Despite this swing, I think three things are important to understand.

One is that Creative Destruction and innovation-based capitalism are *not* something that firms enjoy or necessarily thrive on, despite all the hullaballoo. Entrepreneurs and innovators may thrive, but established and powerful firms often do not. As Schumpeter saw, Creative Destruction creates an environment of paranoid anxiety and, as we will see in our coming readings, myriad difficult challenges for firms wanting to maintain their position. Most established firms, for sure, would prefer *not* having to innovate, but do so because others do it and they therefore have to. For those established firms, innovation is a race to the bottom. Fortunately, that race is what makes *consumers* better off. All that struggle to expand output and reduce prices puts more, better and cheaper goods into the hand of a greater share of consumers.

Another is that ideology can to some extent substitute for *actual* competitive pressure, in the Schumpeterian sense that a pervasive ideology of innovation can make firms more paranoid *and* in the sense that an ideology can make it necessary for firms to innovate in order to stay legitimate in the eyes of relevant stakeholders. A strong ideology can force firms to engage in innovation as a ceremonial, rhetorical practice. This line of thinking would bring you into a somewhat different understanding of the role of innovation in firms, which is undoubtedly a big part of the picture and the subject of much of what is called institutional theory. It is something that falls somewhat outside of the scope of this particular book, but definitely is a stable of courses on organizational theory.

The third thing to understand is that the pendulum can swing again. Just as a change in ideology that nobody planned made the modern world possible, a change in ideology that nobody plans or expects can happen again and undo it. And that certainly would be tragic and dystopian.

1.3. Abernathy: From fluid to specific

Schumpeter made us aware of a dual challenge that capitalist economies present to firms. On the one hand, there is a short-term challenge of competing on price and quality. This is the competition that occurs at the ‘margin of profits’, where firms seek to extend their power over existing markets by offering something slightly better than their competitors. On the other hand, there is the long-term challenge of competing with new ideas that strike at the ‘center of profits’. This is what Abernathy refers to as the challenge of competing *both* on productivity and incremental innovation *and* on major innovation. Why is this a challenge? Because, as Abernathy writes in his book’s opening chapter, “Stated generally, to achieve gains in productivity, there must be attendant losses in innovative capability; or, conversely, the conditions needed for rapid innovative change are much different from those that support high levels of production efficiency” (p. 4). In other words, there is a *trade-off* between competing in the long-term and short-term. To relate this back to Schumpeter, the existence of such a trade-off is the explanation for *why* there is Creative Destruction: firms that successfully pursue production to thrive short-term will not (very likely) survive in long-term, innovation-based competition.

Two things lead Abernathy to this far-reaching argument: An observation of industry difference and an observation of organizational change over time.

The first is a general observation that industries differ in what sort of innovation matters and that there is a relationship between this industry-level characteristic and how firms in those industries *organize*. Most industries are characterized by high-volume production for well-established markets. Primacy is given to a particular production process that is highly optimized and very capital-intensive, and (as a consequence) products are not only standardized but also partially defined by what is possible within the limits afforded by these production processes. In these industries, the kinds of innovation that matter most are process innovations and incremental product innovations. Some industries are very different. In them, major product innovation is what matters and firms tend to be smaller and more flexible. Innovation often comes from new entrepreneurial firms. Products tend to be ‘performance-maximizing rather than cost maximizing” (p. 70), but performance criteria also “are typically vague and poorly understood” (p. 70). The production process tends to be very idiosyncratic. In industries characterized by high-volume production, organizations are highly *specific* or what Burns & Stalker have classically called ‘mechanistic’ in their book “The Management of Innovation” (1961). In industries where major product innovation matters most, organizations are highly *fluid*, or ‘organic’.

The second is Abernathy’s own insight that these differences are *not independent*, but a result of organizations transitioning *over time* between these two states as their industry matures. The change over time from a fluid to a specific is, Abernathy argues, a systematically occurring one and a result of change in the nature of competition in an industry.

Following a major innovation in an industry, organizations tend to be fluid to accommodate the rapid speed of change and the uncertainty of the new technology. Over time, that uncertainty decreases and competition becomes more predictable and more price-focused. As price becomes more central and products more standardizable and therefore standardized (because due to economies of scale, standardization is more profitable than customized, one-off products), organizations become more specific. They start to rely more on standardized production processes and invest in optimizing them by increasing ‘capital depth’ (i.e. using more and more specialized production equipment), but in doing so also sacrifice flexibility. You can think about the difference between a craft shop where everything is custom-made by hand (where efficiency is low but flexibility high) and an assembly line (where efficiency is much greater but flexibility very hard to achieve). Over time, the form of innovation that matters also changes. Initially, what really matters is product innovation and there tends to be a lot of it. This, however, drops over time as the product becomes more standardized. Process innovation, by contrast hardly exists at the outset. As the product becomes standardized, process innovation can increase and with it will come increased efficiency. At a certain point, process innovation will also begin to taper off. All of this is summarized in figure 4.1, which really is the model that is most essential thing to grasp in the chapter.

In contrast to the idea that industries have certain more or less stable characteristics, Abernathy's view is one of firms and productive units being constantly in a process of *transition*, trending away from a fluid state towards a specific one (figure 4.2 adds some depth to the model in 4.1 by describing the transitional stages between the fluid and the specific). A watershed moment in this process is the settling upon a *dominant design*. A dominant design is a more or less standard 'blueprint' for how a product should be designed and function. Dominant designs exist today for most of the products that we surround ourselves with. Laptops, for instance, all follow a pretty standard design: a clamshell, foldable setup with a screen on the upper half and a keyboard in what is called a QWERTY-configuration on the lower half. Cars, smartphones, hearing aids, bicycles and other products similarly follow a standard design blueprint. "The important economic effects of a dominant design afford a degree of enforced product standardization, so that production economies can be sought, and provide a benchmark for functional performance competition, so that effective competition can take place on the basis of cost as well as product performance" (p. 75). This in turn makes possible and necessary a whole series of other changes – in organization, in production process, in marketing, etc.

Abernathy's argument makes it apparent why we need a classification of different forms of innovation. Like Schumpeter, Abernathy recognizes that both product and process innovation matters, but also that they require different conditions, matter relatively more in different competitive situations and seem to more or less crowd each other out. There are also other interplays between different innovation forms as firms transition from fluid to specific forms of organizing. We might imagine, for instance, that organizational innovation might be necessary to make that transition. Scaling up production might bring about a need for new methods of sub-dividing labor and new ways to manage and coordinate workers and units within the organization. We can also imagine that firms may seek to open up new markets over time, both to ensure that the expanded capacity that can come with economies of scale can be fully utilized and because it becomes apparent that larger and larger groups may be ready to consume the product. This raises the question of how these different sorts of innovation are supported by respective fluid, transitional and specific structures.

The larger question that Abernathy leaves us with, however, is *whether the trade-off between efficiency and innovativeness must be a hard one*, whether the dilemma is so grave. This is a question that we will keep returning to throughout the remainder of the book.

1.4. March: The trade-off between exploitation and exploration

In coining exploration and exploitation, this paper introduces what is perhaps *the* most widely used set of terms in current innovation management research, save perhaps for Schumpeter's Creative Destruction. Going forward, we are going to see this set of terms come up again and again as an anchor to many discussions of the challenges of innovation management. Where they do not come up directly, we will be applying them in our interpretations to make sense of many of the source texts for the book and to tie them together. In a very clear way, they make explicit the nature of innovation and get us if not to the heart of the productivity dilemma, then at least closer to it. In this spirit of seeing the terms as helpful tools for understanding and theorizing, I encourage you to dwell relatively more on absorbing them than on the simulation models that March presents us with in the paper's second and third sections.

But what do the terms mean, beyond their immediate intuitive meaning? March is by intention not extremely clear and certainly this is part of the reason the concepts can get used so widely and inspire so much thinking and discussion. But in general terms, "[e]xploration includes things captured by terms such as search, variation, risk taking, experimentation, play, flexibility, discovery, innovation. Exploitation includes such things as refinement, choice, production, efficiency, selection, implementation, execution." (p. 71). Thinking about this in Abernathy's terms, exploration coincides with the emergence of new major innovations and with pre-dominant design fluidity. Exploitation seems to reflect a much more specific form of activity.

Exploitation and exploration refer to different ways of approaching organizational learning. "The essence of exploitation is the refinement and extension of existing competences, technologies, and paradigms" (p. 85), while "[t]he essence of exploration is experimentation with new alternatives." (p. 85). Exploration deliberately has a greater element of engaging with the unknown than exploitation, but both are necessary for organizations that compete in the dynamically evolving circumstances of capitalist economies, because they speak to the dual anxiety that capitalism, in Schumpeter's view, creates for firms. Exploiting without exploring can leave the organization trapped and unable to keep up with changed circumstances. Exploring without exploiting will lead to many ideas, but too little execution to realize the potential benefits. The challenge then is one of *balance* between the two and, for Abernathy, as *fit* with the industry and the competitive dynamics that firms must engage with.

It is worth dwelling on the very different benefits from these activities. The benefits of exploitation, because exploitation tends to be applied and builds on knowledge already well understood, is characterized by "speed, proximity, and clarity" (p. 73). By contrast, "returns from exploration are systematically less certain, more remote in time, and organizationally more distant from the locus of action and adaption" (p. 73). The uncertainty and distant pay-offs of exploration are essential to understand.

Understanding these benefits is essential not only because firms must *balance* the two types of learning in such a way that the organization can actively compete in current conditions, but does not end up trapped in those conditions once change occurs. This challenge is fundamentally Schumpeterian, because the fact of capitalism is that change *will* happen. However, firms must not only engage in both exploration and exploitation. The real challenge is that there is a *strong trade-off* between the two. To say that there is a strong trade-off is to say that the two activities are mutually exclusive, that choosing between them is a zero-sum game, that doing more of more detracts from the other. This is the case between the two activities *compete for scarce resources*. Allocating resources to one means that the other will not receive resources and will not, the hypothesis would go, get pursued.

Precisely because the benefits of exploration are distant and uncertain, they also tend to be less attractive to invest in than exploitation, where benefits materialize sooner and with greater certainty. Much more than exploitation, exploration can be seen as an experiment and as a 'bet' on an untested hypothesis. Exploration will often involve making choices in, to borrow a term from Moore (1991), 'high risk, low data' situations, while exploitation will involve ‘low risk, high data’ situations. The two can be hard to compare, but to the managerial eye the business case for exploitation tends to look substantially better. It will typically make better ‘business sense’ and may even be necessary to compete in the short term. The catch-22 of that is that – to the extent that exploitative projects *do* exclude exploration – it is (put dramatically) a suicidal path long term.

There’s a particularly dramatic image that can invoke the nature of this situation. Consider Ulysses, the hero of Homer’s Odyssey, returning to Ithaca after the Troyan Wars. On his way, his ship must sail through the Strait of Messina. That passage – like the balance between exploitation and exploration – is far from trivial. On one side is Charybdis, a great sea monster, creating whirlpools that will pull entire ships underwater and destroy them. On the other side is Scylla, a nymph turned into a six-headed monster, that will snatch sailors from the ship and devour them. The Strait is narrow and there is hardly any room for error – sail too close to either side, and the results may be dire. On the one hand, there is exploitation, to the exclusion of exploration – a seemingly safe path, but also one that leads to what sometimes gets called a ‘success trap’ (and March himself refers to as ‘suboptimal stable equilibria’). That, however, sets you up for being swallowed by the whirlpool of Creative Destruction. On the other, there is exploration, to the exclusion of exploitation, where you absorb all the downsides of exploring without reaping the benefits. That is what gets called a ‘failure trap’ and that too will destroy you, not in a whirlpool but on the rocks or between the monster’s teeth. What is one to choose?[3]

That March is talking about *choice* is also essential and a difference between him and Abernathy. Due to his focus on organizations trending towards a specific state over time, Abernathy easily comes across as deterministic. Over time, industries will change, competitive dynamics will change and organizations will have to change with them. The arrow of time points away from fluidity towards specificity. For March, it is much less deterministic. Implicitly, “buried in many features of organizational forms and customs, for example, in organizational procedures for accumulating and reducing slack, in search rules and practices, in the ways in which targets are set and changed, and in incentive systems” (p. 71) many choices are made that affect the balance that might be struck between the exploration and exploitation. Seemingly obvious choices to increase efficiency, for instance, might inhibit how much exploration becomes possible. Equally obvious choices to be highly flexible might inhibit how much exploitation is possible. Yet despite such implicit choices, March would think that organizations can also explicitly choose how to prioritize their activities and can make decisions about pursuing exploitative paths or exploratively seeking new opportunities. Bound of course by implicit prior choices and the limits of rationality and certainly disciplined by Abernathy’s competitive dynamics, organizations and individuals in them are in March’s view ‘free’ to explore or exploit.

Dealing with that choice is in different ways going to be at the center of what we will subsequently read. Dealing with it might be a matter of finding an optimum balance, of escaping the constraints of the trade-off, of finding ways to separate exploration from exploitation, of reflecting on implicit choices that bind explicit ones, of overcoming the challenges of uncertainty and distant pay-offs, of understanding the emergent effects of one’s choices on others, of appreciating the many limits to freedom of choice. But the nature of capitalist economies is that firms *must* deal with the choice.

[3] At the advice of the sorceress Circe, who in some accounts was the one who transformed Scylla from a nymph to a sea monster, Ulysses chose to steer close to Scylla. While taking that path, he was momentarily distracted by Charybdis and lost six men to Scylla. His ship made it through.

Chapter II:
Technological improvement and industrial change

Texts in this chapter:

Foster, R. 1986. **Innovation: The attacker's advantage.** Summit Books.

Anderson, P. & Tushman, M. L. 1990. Technological discontinuities and dominant designs: A cyclical model of technological change. **Administrative Science Quarterly**. 35(4), 604-633.

Vincenti, W. G. 1994. The retractable airplane landing gear and the Northrop "anomaly": variation, selection and the shaping of technology. **Technology and culture.** 35(1), 1-33.

Adner, R. & Kapoor, R. 2016. Innovation eco-systems and the pace of substitution: Re-examining technology S-curves. **Strategic Management Journal.** 37(4), 625-648.

In this chapter, we are going focus on technology and how technological change impacts industries. The central metaphor for understanding technological development is the technology 'S-curve'. The central way to understand new technologies' impact on industries is as cyclical, as short periods of rapid and dramatic change followed by long periods of incremental improvement. We will consider both and explore how they connect to one another, to the dynamics of Creative Destruction, to the dominant organization of industrial firms and to the balance required between exploration and exploitation.

With these basics in place, we are going to complicate the relatively straight-forward story they produce. We are going to examine how technological performance at the system level is a function of sometimes highly unpredictable changes at the sub-component level and sometimes of changes at the ecosystem level. Both can render predictions of technological change and competitive dynamics more difficult and open up new avenues for innovation management action.

2.1. Foster: The technology S-curve

Technologies do not come into the world with their full potential performance realized or even clear. On the contrary, technologies have a life cycle of being introduced, maturing and finally stagnating and becoming obsolete as new technologies are introduced. Over the course of that life cycle, the performance of a technology is going to change. Foster's first big idea in his book is that a technology's performance is going to change in predictable ways. He writes: "Initially, as funds are put into developing a new product or process, progress is very slow. Then all hell breaks loose as the key knowledge necessary to make advances is put in place. Finally, as more dollars are put into the development of a product or process, it becomes more and more difficult and expensive to make technical progress. Ships don't sail much faster, cash registers don't work much better, and clothes don't get much cleaner." (p. 31-32). In other words, the rate of technological improvement tends to be slow when the technology emerges, then more rapid as maturity is reached and then slow again before grinding to a virtual halt. What drives improvement is *investment* and if you plot a technology's performance as a function of investment, the line you get is basically a drawn out 's' (see the curve on p. 31).

We can think about this through the concepts provided by March and Abernathy. Initially, a technology is *not* well understood and a lot of effort is needed to figure it out. Developing it is inherently and necessarily explorative. Organizations working with the technology will tend be fluid, essentially because there is nothing to be specific around. If the technology is poorly understood, it is hard to know *how* to organize a specific production process and what to optimize on. As investment builds up and more experiments have been conducted, exploration can give way to exploitation, especially when a dominant design emerges and sets a direction for future innovative efforts. When the industry closes around a dominant design, what Italian economist Giovanni Dosi calls a 'technological paradigm' (1982) emerges and sets a trajectory for future development work. The dominant design defines a set of questions and problems to work on in order to improve the technology and a set of approaches for doing that work.

Towards the end of the life cycle, Abernathy would depict organizations as becoming so specific as to be almost incapable of innovation – we are beyond the point where process innovation simply crowds out product innovation, and instead in a state where efficiency simply crowds out innovation generally. March would surely think of the technology life cycle as following a pattern of, initially, 'pure' exploration in the interest of discovery and then over time trending towards 'pure' exploitation in the interest of efficiency and predictability and reduction of variance, with various weightings of the two learning activities in between. Foster has a somewhat different perspective in advancing the idea that technologies have *limits* at the top of the S-curve where there simply is no more innovation that can viably be done, because the cost of further improvement becomes so high. Where Abernathy would think of innovation as an effect of organizational conditions and March as a one of choices and resource allocation, Foster would think of technologies as having inherent 'natural' potentials that can get exhausted, much in the same way that a well can run dry. As technologies mature and reach their limit, ideas for improving them become harder and more costly to find. That is point at which ships do not sail much faster, no matter how many masts you put on them, because physics simply constrain further improvement (for other technological paradigms, it may be other laws of natural or social science that constrain improvement).

To see that ideas indeed do get harder to find as technologies mature[4], we can look at some of the recent evidence from Nicholas Bloom and three of his colleagues. In their paper, they look at *research productivity*, which is basically Foster's improvement rate and thus an expression of how much investment it takes to sustain constant exponential growth. In semiconductors, where performance has improved roughly 35 per cent every year almost since Gordon Moore in 1965 made the prediction that would become Moore's Law, research productivity is dramatically falling. Today, there are 18 times more researchers working on developing semiconductor technology than there were in 1971. A similar pattern is evident in agriculture, where crop yields (per unit of land) have roughly doubled since 1960 with relatively constant growth rates during the period.

[4] The authors imply that improvement *in general* is getting harder to achieve, which, to my mind, is not at all what their data show. Rather, they look at technologies that have been subject to intense investment over long periods, have matured quite far and *therefore* now face very low returns for further investment. Others have made the argument that improvement in general is getting harder, such as Robert Gordon (2016), but that argument has been quite convincingly rejected by e.g. Erik Brynjolfsson and Andrew McAfee (2014). The main difference between optimists and pessimists, and the reason that I think both Bloom and his colleagues and Gordon are over-dramatizing, is that they do not take account of *new* technologies that are *early* in their lifecycle, i.e. are not Schumpeterian enough as I will expound on later.

However, it takes more than 20 times the effort that it took in the 1960's to sustain that growth rate. The pattern is less clear, but similar in medicine, where we also see that new drugs and treatments become increasingly expensive and have a decreasing impact on health outcomes. The dystopian reading of this is that we will 'run out' of improvement because with decreasing productivity we will also run out of researchers. A Schumpeterian perspective would be much more optimistic, assuming that new technologies come along and, being earlier in their life cycles, offer opportunities for less costly improvement. This is what Foster calls 'technological discontinuities'.

There are a few things to note before we turn to that topic, however. Do note that not all technologies follow the S-curve pattern. If you look at Foster's account of the development in tire materials (p. 124), you will see several technologies with quite different development trajectories. Note also that it is a common misconception to depict S-curves with *time* on the x-axis. When you depict improvement as an effect of time, you (inadvertently, perhaps) assume that improvement 'just' happens and you end up obscuring the fact that improvement is *very much* an effect of investment and effort and that improvement cannot be taken for granted. Certainly, if investments were constant over time, this would not be much of a conceptual problem. Alas, investment in technology sometimes follow a boom-and-bust pattern of inflated expectations and shattered hopes as investors flock to and rush away from overhyped and spectacularly failing technologies, while in other cases investment can be highly directed and coordinated across different actors in an industry.

The final thing to note about technology S-curves at this point is that both the phase of rapid improvement and the eventual limit where improvement becomes hard to sustain often appear clear in retrospect, while forecasts about technology development are inherently shrouded in uncertainty. The primary source of uncertainty is where the *limit* of a technology lies. The limit of improvement in semiconductors, for instance, has been continuously debated, with repeated declarations that 'Moore's Law is dead'. Similar discussions surround many other technologies. It turns out to be hard to predict, because the limit is not only a function of the technologies inherent, 'natural' properties, but also a function of human ingenuity. Moreover, there is an element of choice that contributes to *shaping* the S-curve. In their book on entrepreneurial strategy, Scott Stern and Joshua Gans (2018) argue that entrepreneurs when choosing to innovate on a technology can *choose their S-curve* by strategically pursuing more or less explorative trajectories. An exploitative strategy might involve quickly implementing a technology, focusing more on developing it from its current level to a particular application and thus settling for a comparatively low ultimate limit. An explorative strategy might involve a broader, longer and costlier search before settling into on a technological paradigm, thus providing the foundations for an S-curve with a higher ultimate limit. In this way, uncertainty also becomes a function of other market actors' investment decisions, with those decisions at the micro-level creating an aggregate performance improvement, defined by the rate and direction of inventive activity. There are also other sources of uncertainty that we will return to in subsequent chapters.

The idea of choosing an S-curve can seem somewhat opaque, but consider the very current example of self-driving cars. While this technology is imminently close to realization, it is also clear why it has been a very difficult problem to solve. A vehicle capable of autonomous driving at high speeds in complex and erratic environments and of responding safely to that environment requires technology to have reached a very high level on the S-curve to be at all viable. For that reason, extensive exploration and substantial investment has been necessary, just to get *close* to realizing what could be a 'minimally viable product'. Now contrast this with what could *also* have happened. Instead of aiming for a fully autonomous *car*, producers could have looked at the same basic technology and taken a much more exploitative route, focused on getting a product to market much faster, if they had instead aimed for an autonomous package-delivery robot. Such a robot could suffice with driving at substantially lower speeds (say, 4 km/h instead of 120 km/h), using the sidewalk instead of the highway. Instead of having to deal with high-speed collision avoidance, it could simply stop if it encountered a situation that the software was unable to handle. Without having passengers on board, safety demands would also be much lower for this robotic mailman than for the autonomous car. It would, in other words, have represented the pursuit of a much less demanding minimal solution, a much more exploitative S-curve and possibly resulted in a much lower ultimate limit to self-driving technology, because foundational and difficult-to-achieve insights might not have occurred, because they would not be necessary.

Foster's second big idea is of 'technological discontinuities'. A technological discontinuity refers to "a break between the S-curves" (p. 35), where a new technology replaces an old one. Typically (but not always), as one technology S-curve reaches it limit and begins to be subject to decreasing returns on innovation investment, new technologies emerge. Initially, such new technologies tend to exhibit weaker performance than old ones. It is only over time that they become superior, because they start to mature and rapidly improve at a point where the old technology has effectively reached its limit (see p. 102 for a graphic presentation). The first digital cameras were by most measures not as good as cameras using photographic film, but they improved more rapidly because photographic film had exhausted most of its improvement technology. "If you are at the limit, no matter how hard you try you cannot make progress" (p. 34), and that will set you back in competition with new solutions that *do* progress.

Technological discontinuities "are almost always brutal" established firms, because when a new S-curve emerges its technology tends to *not* rely on "the same knowledge that underlays the old one but [on] an entirely new and different knowledge base. For example, the switch from vacuum tubes to semiconductors, the switch from propeller-driven planes to jets, the shift from cloth to paper diapers, the switch from records to tapes to compact discs…" (p. 35-36). Each of these represent a paradigmatically different way of thinking about problems and solutions respectively for controlling electricity flows in computers, for air plane propulsion, for water absorption and for storage.

This leads Foster to his third central idea, namely that when technological discontinuities arise, they tend to favor new entrants to an industry and to challenge established firms. Attackers, he suggests, tend to be afforded advantages by technological discontinuities. When "attacks are launched, they are often unnoticed by the leader, hidden from view by conventional economic analysis. When the youthful attacker is strong he is quite prepared for battle by virtue of success and training in niche markets. The defender, lulled by the security of strong economic performance for a longtime and by conventional management wisdom that encourages him to stay his course, and buoyed by faith in evolutionary change finds it is too late to respond. The final battle is swift and the leader loses. His attempts to defend his employees and shareholders fail. Doomed by doing too little, too late" (p. 37).

There are lots of moving parts in that description. We will be returning to the role of niche markets when dealing with Clayton Christensen's ideas of Disruptive Innovation and the inertia of incumbents with Henderson & Clark's ideas of Architectural Innovation (both in chapter 3). For now, consider the way this statement reflects Schumpeter's depiction of Creative Destruction. Essentially, Foster's technological discontinuities are the sort of innovative step change that strikes at the heart of profits and bring about the dynamism of Schumpeterian capitalism. For this reason, one of Dosi's observations about new technological paradigms is that they bring about the emergence of ‘Schumpeterian firms’. The hypothesis here is that attackers, to use March’s term, can pursue a purely explorative strategy and are unconstrained by having to negotiate a balance between exploration and exploitation and that this freedom for some attackers will outweigh the liabilities of competing against large, established and efficient organizations with ample resources. In Abernathy’s terms, they also do not have to undo a highly specific organization of their operations, but can begin from an entirely fluid and more dynamic starting point, the assumption being that such specific organizations *are* difficult to shift.

Taken together, Foster makes us alert to the central importance that current and ultimate performance levels *and* rates of technological improvement play in determining competitive dynamics and in spurring Creative Destruction. We will return to this in the subsequent chapter. But perhaps more than Foster makes clear, forecasting of these things is inherently speculative and irreducibly uncertain. Performance and improvement are both highly complex phenomena, with multiple determinants within and beyond the focal technology. This will be a continuing theme throughout the book. What is clear, already at this point, is that pursuing innovation and exploration should basically be understood as experiments, as tests of an ex ante unproven hypothesis about how technology and numerous other related factors will develop. Under certain conditions, some organizations will be better at those experiments than others and some, if not most, of those experiments will fail.

2.2. Tushman & Anderson: Technology cycles

There is a circle to be squared in reconciling two ‘pictures’ that we have so far encountered of innovation. One image that we have met is Schumpeter’s, the other Abernathy’s. For Schumpeter, the capitalist economy is in a state of perpetual turmoil. The perennial gale of Creative Destruction constantly blows, spurred on by innovation. Innovation is the stuff that competition is made of and the result is constant change. This is the ideology of innovation at full throttle. For Abernathy, the industrial organization is perpetually seeking stabilization, specificity and regularity. The search for efficiency drives out innovation, first by seeking a dominant design and then investing in capital deepening to the point where innovation can no longer occur. Efficiency achieved through economies of scale is the stuff that competition is truly made of and the result is a roadblock to innovation and ultimately stagnation. This is, if anything, the inverse of the innovation ideology. As it turns out, Foster’s idea of technological discontinuities is a good place to begin the squaring, because it can reconcile the two opposing views into a model of ‘punctuated equilibrium’. There is, generally, stability. And then comes the Schumpeterian shock of paradigm-changing technological discontinuity.

This idea of innovation driving change through punctuated equilibrium is what Anderson & Tushman develop with their cyclical model. This has emerged as the canonical way to think about how technological change brings about industrial change. The model (depicted in their paper’s Figure 1) presents industries as transitioning between two states. Transitions between these states are driven by particular events. Eras of Incremental Change are the state that industries are in most of the time. These are long (Abernathian) periods where firms pursue increased scale and efficiency through process innovation. The technological paradigm underlying the industry is mostly settled and development is focused on marginal improvements. This state ends with the introduction of a (Fosterian) technological discontinuity, where a new technological alternative emerges to challenge the technological paradigm. The technological discontinuity brings the industry into a (Schumpeterian) Era of Ferment, where competition is driven by innovation, intense and rife with uncertainty. The decisive event that lets the industry settle back into an Era of Incremental Change is the emergence of a dominant design. This state prevails until another discontinuity emerges, and thus the cycle continues. This is a model is long periods of evolution, interrupted by short spurts of revolutionary change. That is punctuated equilibrium.

Reading this model through March's terminology, it is apparent that Eras of Ferment are characterized by exploration. Eras of Incremental Change, by contrast, are characterized by exploitation. The emergence of a dominant design allows firms to transition from exploring to pursuing the more certain and more proximate gains from exploitation. The exploitative nature of eras of incremental change, is evident in how competition mostly takes the form of elaborating the dominant design. That competition, however, can still be very vigorous as made clear by for instance the Cola Wars the 1980's where Pepsi and Coca-Cola struggled for dominance of a well-established market, but competed largely *without* substantial innovation. Colored sugar water did not change much. What changed was essentially the means and intensity of persuasion: advertising, packaging, endorsements and the like.

In eras of ferment, two forms of competition play out and innovation and exploration play key roles in both of them. The sort of competition that Anderson & Tushman, building on Foster, describe as *substitution* is competition between old and new technologies. Generally, old technologies do not go quietly into obsolescence, but rather compete vigorously with new ones. A classic examples of this might be the competition between IBM and Remington Rand when electronic computers based on vacuum tubes were first introduced to replace computer based on switches. A more recent one might be the contest between Blockbuster and Netflix, when Netflix introduced DVD rental by mail to challenge physical stores, described in Keating (2012).

A less widely appreciated sort of competition is the competition that occurs between different instantiations of the technological discontinuity, what they term design competition. The classic example here is the contest in the videotape market that played out between VHS and BetaMax, beginning in the 1970s, described by Cusumano et al (1992). Sony launched BetaMax in 1975, JVC launched VHS in 1976 and through to the middle of the 1980s an intense contest played out to become the standard format, i.e. the dominant design. A more recent example might be the contest between Blu-Ray and high definition DVD or the contest of Large Format Cinema between, most famously, IMAX and Dolby but also several others. An older one, described in detail in David (1985), is the contest to become the standard typewriter keyboard layout, which QWERTY won.

It is an interesting fact to note that in Anderson & Tushman's analysis, it is never the discontinuity that becomes the dominant design. This tells us something about the dynamics of design competition in Eras of Ferment and the advantages of being first to market that we will return to.

It is the model that most researchers focus on as the central take-away from the paper, but the underlying study is really admirable in its detail of following three industries over such long periods as to observe several discontinuities in each. While not quite as central as the circular model itself, there are several points to take with you going forward that deepen our understanding of technological discontinuities and add to our repertoire of theoretical distinctions.

For one, the kind of discontinuity that Anderson & Tushman observe differs from Foster's in an important respect. Where for Foster a new technology starts off with performing that is inferior to the old, Anderson & Tushman's discontinuities are characterized by performing *better* at the outset. The visual representation of this appears in Figures 2a and 2b, where the discontinuities appear as sudden bursts of increased performance. This is noteworthy not only for reminding us that Foster's pattern is a stylized one that not all discontinuities follow, but also for interpreting Anderson & Tushman's findings. It appears uncontroversial that this has to lead to considerably less uncertainty at the emergence of a discontinuity than we would otherwise expect. We could, with March, expect that to make exploration of new opportunities appear less risky, and hence more attractive for incumbents than would be the case of discontinuities where the new technology is initially inferior.

Secondly, they alert us to the possibility that all discontinuities are not created equally. Where Foster sees discontinuities as based on new knowledge, Anderson & Tushman argue that a discontinuity can be both competence-enhancing and competence-destroying[5] and that this matters for how an era of ferment plays out, essentially by changing the relative advantages of incumbents over new entrants. This is the case, because competence-enhancing technological discontinuities essentially build on the knowledge that already underlies existing technologies, while only competence-destroying discontinuities fit Foster's model of relying on new knowledge. A discontinuity, for Anderson & Tushman does not imply a technological paradigm shift, but only a dramatic change in performance levels. That conceptual clarity can be worthwhile to bear in mind, because we can see that we suddenly have two meanings related to the same term and can expect two different causal mechanisms to unfold around them.

Third, technological substitution is *not* guaranteed, again in contrast to Foster's depiction. Because the emergence of a new technology can lead to increased investment in an old one, the opportunity exists for a new technology to not improve sufficiently and sufficiently quickly to supplant and old one, even as the old on is late on its S-curve. When a discontinuity is competence-destroying, we can expect incumbents to invest heavily in competing against a new technology to keep it from gaining a foothold. This is the case, because making the shift to that new technology will be so relatively difficult for them and, we can assume, so relatively unlikely to be successful.

Fourth, what becomes the dominant design is *not* a matter of technological superiority. Rather, it is a quite complicated matter in which many elements play a part: choices between which customers to pursue most aggressively, strategic maneuvers in building alliances and supply chains, evolving customer perceptions, etc. It is a question of dominant designs *emerging* as outcomes of directed human action, but not as outcomes of any particular person's intention or any particular firm's strategy. That any person or firm could actively steer the outcome is one of those interpretive mistakes that are easy to make in superficial hindsight.

Finally, the model has an alternative outcome that is not developed in the paper. It posits that an Era of Ferment ends with the emergence of a single dominant design and that the vanquished technology fades into obsolescence. Neither of those are necessarily true. An Era of Ferment could viably result in several dominant designs and an ensuing splintering of an industry into several sub-industries. A vanquished technology could

[5] They originally developed that argument in a prior paper, titled 'Technological Discontinuities and organizational environments' and published in 1986.

also 're-emerge' after having been judged obsolete. Raffaeli's work on the Swiss mechanical watch industry (forthcoming) is exemplary of both dynamics. In the 1980's, mechanical watch technology was widely regarded as obsolete due to the introduction of battery-powered quartz-based digital watches – a truly competence-destroying technological discontinuity. However, mechanical watches reemerged in the 2000s as a distinct market within the watch industry that largely does not compete with quartz watches. That possibility, while perhaps not frequent, is worth bearing in mind.

2.3. Vincenti: Technological interdependencies and performance shifts

This paper provides a detailed analysis of a particular moment in aviation history that tells us something quite profound about technological discontinuities and eras of ferment. More specifically, it provides an account of how retractable landing gear replaced fixed landing gear as a part of the dominant design for airframes in the 1930s and does so by studying in details the seemingly anomalous design choices made by one particular airplane designer – John Northrop. If that seems supremely obscure to you at the outset and if the details of the study seem excruciating, you could be excused for thinking so. But in that obscurity and in all those details lie, I think, some great lessons for especially management students and scholars, used as we are to more sweeping accounts of technological change, its strategic implications and the visions of great men. For good measure, I can assure you that it will not be the last time that this book picks up papers like it.

The paper has its own theoretical ambition, which is to counter what Vincenti calls 'the usual view' that the transition to retractable airplane landing gear was "an essentially reasoned and ordered affair" (p. 2) and that "the desirability of the retractable gear was obvious and subject to simple and straightforward engineering assessment" (p. 4). Instead, Vincenti wants to advance the biological metaphor, drawn from evolutionary theory, of blind variation and selective retention (elaborated from p. 22 and forward). To be honest, I actually find that part of the paper to be the least interesting – many other parts are more interesting.

Read 'just' as literary fiction, the historical analysis conveys something fundamentally important about eras of ferment. What it conveys so well, I think, is the incredible uncertainty that surrounds newly emerged technologies. Having read March, this would not surprise us, but Vincenti's description is interesting for how it helps us understand how an otherwise highly esteemed and talented airplane designer finds it so difficult to navigate between technological alternatives and why he is so late to transition to what *in hindsight* appears to be so obviously superior a technology. Especially when it comes to technology, this study shows us how foresight is infinitely harder than hindsight.

This caution in itself would be reason to read the paper. We can also learn something about exactly what it is about technology and eras of ferment that make uncertainty so inherent, though. Consider, for instance, how Northrop must evaluate the relative merits of retractable landing gear over fixed landing gear. Contrary to Foster and to Anderson & Tushman who conceive of performance as mono-dimensional, Northrop must deal with *multiple possible performance dimensions*. If we see landing gear as a *sub-system* in the larger system of an airplane, which also consists of other sub-systems (e.g. engines, wings, steering system, etc.) it is clear that the decision between retractable and fixed landing gear is a complex one with trade-offs between at least five important requirements: The sub-system must be lightweight so to not weigh down the airplane; it must be aerodynamic to avoid drag; it must be reliable to ensure safety; it must be easy to maintain; and it must be seen as competitively priced relative to alternatives.

In some cases, the balance of those trade-offs will crucially depend on technological developments *in other sub-systems*. In Northrop's case, the importance of aerodynamics *increased* over time as engines improved and airplanes consequently flew faster. That engines would improve was not obvious, but when they did the disadvantages of fixed landing gear grew relative to retractable landing gear that folds away inside the air-frame. As Anderson & Tushman show, sometimes old technologies get improved when a new technology gets introduced. This was the case for fixed landing gear, where the introduction of 'trouser' covers for the landing gear (pictured in figure 4) made them appear more aerodynamic than retractable gear, but only for a certain time. When engines further improved, that advantage went away.

In other cases, change in one sub-system will compromise the performance of other sub-systems. In Northrop's case, he had developed a wing design that was for all intents and purposes revolutionary in his time. If he had implemented retractable landing gear, he would have compromised with the performance of wings, because those two sub-systems were made interdependent by competing for space within the air-frame. The appearance of a trouser design, i.e. relying on an elaborated version of an old technology, made it seem entirely reasonable to stay on an old S-curve for the landing gear sub-system to maintain a leading technology in the wing sub-system.

Finally, the performance of a sub-system on a particular dimension can be suddenly jolted by inventions made in very different fields. Where Foster sees technological improvement as smooth, what Vincenti's account shows us is that retractable airplane landing gear was *suddenly* made dramatically more viable. Early versions of retractable landing gear were raised and lowered by hand, which can be somewhat problematic for a pilot who has plenty of other things to do during take-off and landing and thus reduce the desirability of moving to that new technology. A later version was hydraulic, but this relied on a wholly unreliable and maintenance-demanding system of sealing fluid within this sub-sub-system. Only when the O-ring was invented as part of developing a braking system for streetcars and later implemented in the hydraulic system, did retractable airplane landing gear as a sub-system become really viable, because it made the sub-sub-system that raised and lowered landing wheels more reliable and less maintenance-intensive. "Thus did a seemingly minor component play an important role in the success of a much larger and more visible device." (p. 20).

These kinds of considerations, while seemingly obscure and far outside the domain of business and management, are crucially important for innovation and technological substitution. What is instructive about them is that "This kind of situation may be present invisibly in many cases." (p. 20), and by virtue of their invisibility be disregarded as trivial in hindsight. What we can take away from reading about them here is the understanding that innovation is, in Joel Mokyr's (1990) terms, hopelessly overdetermined: there are *so many* things that matter. In business and management research, we tend to look for those variables that are most important in explaining a phenomenon in order to get 'clear' stories and prescriptions that apply across contexts. But when we immerse ourselves in the particularities of practice, it becomes clear that there are many, many dynamics that interact in order to produce particular, emergent end-results.

2.4. Adner & Kapoor: Eco-systems and the pace of substitution

When a technological discontinuity occurs, Foster would expect it to bring about technological substitution and that the substitution would quick. Recall the statement about what happens when attackers working on a new S-curve enter a market dominated by old technology: “The final battle is swift and the leader loses” (Foster, 1986, p. 37). To that, Tushman and Anderson’s analysis adds that the question of whether substitution does result from an era of ferment depends on the competitive response of incumbents (e.g. do they invest heavily in elaborating the old technology) and on whether the discontinuity is competence enhancing or competence destroying. Vincenti would add that substitution depends on how the process of technological improvement unfolds, which in turn depends on developments at the level of sub-systems and components and to no small extent on unforeseeable events (e.g. the O-ring). Adner & Kapoor provide a way of thinking about the process of substitution that can systematize this. Rather than think of whether substitution happens, they focus on its *pace* and in doing so reconcile the three other views. Substitution, to them, is bound to happen eventually and so the inherently Schumpeterian question is ‘just’ one of when, how fast and, by implication, how destructive for incumbents.

For them, the question is one of whether substitution begins immediately after the discontinuity emerges and, once it begins, whether it is swift or sluggish. That is the outcome they try to make sense of. To explain the outcome, they rely on two groups of concepts. The first is of an eco-system of complements around a focal technology. A complement is a technology that increased the performance of *another* technology. A landing strip is a complement to an airplane, a light is a complement to a bicycle and, in Adner & Kapoor’s case, masks and resists are complements to lithography tools in the manufacturer of semiconductors. By looking at technologies *together* with their complements instead of at individual technologies, Adner & Kapoor add to Foster’s understanding of technological improvement. The performance of a technology can improve without investment in that technology, provided investment in the ecosystem of the technology’s complements yield improvement. Electric cars, for instance, can improve their range either by advances in solid state battery technology (a discontinuity) or by advances in charging infrastructure (a complement). This leads to the distinction between *performance-as-developed* and *performance-as-used.* What ultimately matters, of course, is performance-as-used.

The second set of concepts are the extension opportunities for old technologies and the emergence challenges of new ones. Extension opportunities refer to the opportunities to improve the performance-as-used of an old technology. This could happen either by improvement in performance-as-developed through costly investments, squeezing the very last improvement out of an old technology. Of more interest to Adner & Kapoor, it could also happen by investment in improving complements. Emergence challenges are the analog of this for new technologies: how challenging is it to improve a new technology’s performance-as-developed and its attendant ecosystem to the point where performance-as-used exceeds old technology ecosystems?

If we distinguish between high and low emergence challenges and high and low extension opportunities and say that all cases of technological discontinuities will face some combination of those, we can create a two-by-two matrix. Doing so is a stable of management research and sometimes it makes perfect sense and at other times it is a ridiculous and unhelpful complexity reduction. In this case, it seems to be a meaningful distinction and it yields the framework in the paper’s figure 4, with four ideal types of substitution.

When we have low extension opportunities and low emergence challenges (quadrant 1), we would expect a ‘Creative Destruction’-type of substitution that is swift and decisive, much like what Foster would imagine. This is the case because there is both very little to do to improve the old technology and little to stop the new one and so superiority is achieved quickly. If, *in contrast to Vincenti’s argument*, we assume that superiority on a single dimension is what matters, market dominance quickly shifts, rendering the old technology obsolete. The opposite would be the case for a situation of high extension opportunities and high emergence challenges (quadrant 4). The old technology can be improved considerably through complements and the new struggles to get off the ground, meaning that substitution does happen, but slowly. In that case, the old technology is ‘robustly resilient’ to competition. With high extension opportunities and low emergence challenges, we see some real competition between the two technology ecosystems and a drawn-out era of ferment with ‘robust coexistence’ (quadrant 2). Rank-ordered, substitution would be fast in situations like quadrant 1, intermediate in quadrant 2 and slow in quadrant 4.

In situations like quadrant 3, where the emergence challenges for the new technology are high and extension opportunities low, incumbents can easily have the initial impression that a new technology is not a threat because it appears so hard to develop to a competitive performance-as-used, thus falling prey to a ‘Resilience illusion’. When we assume that the emergence challenge is not so great as to be unviable, the result of this would be that the new technology enters a market, but only achieves a trivial role until the point where the emergence challenge is overcome. Assuming that this challenge is indeed overcome in one decisive step and not in stages and that it is overcome before an even newer discontinuity emerges, the new technology then very rapidly transitions to market dominance. The resilience illusion is then painfully broken for incumbents, as what resembles creative destruction unfolds.

All this theory is developed up front in the paper and then studied in the context of semiconductor production, focusing on successive generations of the lithography tools that together with two complements are central to that production and largely confirming the model’s hypotheses. Like Vincenti’s paper, the industrial details can appear somewhat obscure, even if Adner & Kapoor trade intricacy of historical detail for a much more fully developed theory. That obscurity, however, is instructive because it is precisely in sensitivity to the details that viable strategies depart from unviable ones, that serious analyses and observations depart from those that just restate ideology. In that sense, what we should take from both Vincenti and Adner & Kapoor is an appreciation for just how much context – in this case technological context – matters for how we apply our theories, and for why having concepts that lets us distinguish in finer detail should is valuable.

In this case, you should take very seriously the idea that performance depends not only of a focal technology, but of an ecosystem of complementary technologies, when you think about technological improvement. Just as Vincenti brought our attention to the possibility that improvement at the level of technological systems can depend on *inside* improvement in sub-systems and components, Adner & Kapoor direct our attention to improvement *outside* the technology. In this way, we can add both a micro-perspective and a macro-perspective to Foster’s model and to our thinking about discontinuities. This will, in turn, shape how Eras of Ferment play out and what consequences they will have for both new entrants and incumbents. You should also take seriously the nuance of distinguishing between different levels of emergence challenge for new technologies and extension opportunities for old ones. Most specifically, the ‘Resilience illusion’ situation is one that we will return to in subsequent chapters as an overlooked ingredient in some of the most dramatic instances of Creative Destruction.

Chapter III:
Varieties of innovation and the nature of opportunity

Texts in this chapter:

Christensen, C. 1997. **The Innovator's Dilemma: When new technologies cause great firms to fail.** (Introduction) Harvard Business Review Press.

Henderson, R. M. & Clark, K. B. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. **Administrative Science Quarterly.** 35(1), 9-30.

David, P. A. 1990. The dynamo and the computer: An historical perspective on the modern productivity dilemma. **The American Economic Review.** 80(2), 355-361.

Shane, S. 2010. Prior knowledge and the discovery of entrepreneurial opportunities. **Organization Science.** 11(4), 448-469.

In this chapter, we are going to look at different forms of innovation, but not along the typical axes of difference. In previous chapters, we have distinguished between innovation in products and services, processes and organizational practices and between innovations that are mostly exploitative and incremental and those that are explorative and radical. The ones we will cover in this chapter are characterized by additional, complicating features that lead them to have different opportunities to new entrants, different challenges for incumbents and different effects on industries. These features include particular aspects of marketing, organization and uncertainty.

We are going to close the chapter by taking a slightly different angle on the relationship between technology and industries. Where we have so far taken that relationship as given, we are going to look at how technologies that are *potentially* general in nature get applied to several different industries through a process of opportunity discovery that can in no reasonable way be taken for granted and what that might imply for technological improvement.

3.1. Christensen: Disruption and unattractive markets

Disruption is perhaps the most ideological idea to float around in innovation management practice and – to a lesser, but non-trivial extent – research. For some reason, it seems to have become the de facto innovation strategy for entrepreneurs, the greatest fear for incumbents and a clear and present danger for all industries. You could be excused for coming into a business school course on innovation management or entrepreneurship thinking that disruption was *all* that mattered and in shamefully many cases you could walk out of them with that belief very much intact. If you read the more hyped-up business press, you would probably find your conviction strengthened. No sector is safe, it appears, from the tidal wave of disruption.

Very few academic theories have this kind of impact and while it is truly fascinating to understand just why disruption has risen to this prominence in the popular imagination, that falls somewhat beyond the scope of this book. Suffice to say that you need to know what people talk about when they talk about disruption, but also to know that there are pretty wide gaps between the theory and how it gets talked about, between its empirical basis and its theoretical argument, and between the claimed and actual scope of disruption.

Of all the ideas that this book presents, Christensen's idea of disruptive innovation – he does not talk a lot about disruption, but rather disruptive change, disruptive technologies and disruptive innovation – is perhaps the most Schumpeterian of all. Schumpeter made us aware that innovation is the great harbinger of industrial change, that creative destruction is an ever-present threat for established firms. Christensen pushes this even further, focusing on "*good* companies – the kinds that many managers have admired and tried to emulate, the companies known for their abilities to innovate and execute" (p. xiii, italics in original). Christensen's argument is about how the very best companies get destroyed *by that greatness*. It is not because of the hubris of greatness, even if the story of a 'fat and happy' company getting toppled makes for an excellent story, but because the very things that make a company great also make it susceptible to being swept along by creative destruction. In that sense, Christensen turns up the anxiety of capitalism that Schumpeter identified to even higher levels, but also tries to show the particular pattern that leads to this kind of destruction – there is, to his mind, a very particular recipe for it.

Disruptive innovation, in Christensen's conceptualization and in marked contrast to how disruption gets talked about, means something *very* specific. What Christensen describes as his 'failure framework' is a recipe in four parts with one result[6]. The first part is that incumbents focus on a *sustaining* innovation trajectory. Along that trajectory, you can find all kinds of technological change: it can be both incremental and radical, it can concern products and processes, it can be competence-destroying or competence-enhancing, it can be discontinuous or not. "What all sustaining technologies have in common is that they improve the performance of established products, along the dimensions of performance that mainstream customers in major markets have historically valued" (p. xix). The second part is that sustaining innovation trajectories *overshoot* market demand. Staying on that trajectory, established firms "give customers more than they need or ultimately are willing to pay for" (p. xx). The third part is the emergence of disruptive technologies. These technologies are often initially inferior to the incumbent technology and underperform them and mainstream market demands. These technologies *cannot* satisfy mainstream customers, but because sustaining trajectories have pushed incumbents 'up market', there is space at the lower end of the market where these new technologies *can* satisfy demand. This is especially the case when new technologies can be used to address markets that are currently *unserved* (i.e. the market does not really exist, because consumers are not currently buying, much like there was no market for PCs during the time of mainframes). The fourth part is the disruptive technologies make for poor investments for incumbents, for Marchian reasons. For one, the development trajectory for the new technology is uncertain and the current performance certainly low. It is clear that it is inferior. Secondly, the market that it serves is certainly unattractive, because the market is small and the margins are low. How the market may develop is highly uncertain. When technology improves, these elements (sustaining innovation, market overshoot, a technological change, unattractive markets) come together in a particular way. Suddenly, the new technology rapidly increases in performance (possibly along a new performance dimension) and becomes relevant for a much larger share of the market, thus pushing incumbents into a progressively smaller market niche. Incumbents are unable to react before it is too late, because "the very

[6] You can look at figure I.1 for a partial illustration of the technological and demand parts of this argument. Do not, however, focus too much on the fact that the figure presents developments as linear. That is, for all intents and purposes, just a more pedagogical way to represent what is more likely exponential change at different rates.

processes and values that constitute an organization’s capabilities in one context, define its disabilities in another context” (p. xxvii). The result is the creative destruction of the incumbent. *That* is what disruption means *in a technical sense*.

It is worth knowing that Christensen’s theory has been challenged quite convincingly in recent years after some twenty years with very limited critical reception. Historian Jill Lepore (2014) has argued that disruption is overhyped as a theory and woefully poorly grounded in data. She contends that Christensens failure framework fails to explain the development of the disk drive industry on which it is actually based. King & Baatartogokh (2015) have gone even further. Examining in detail 77 examples of innovations that are held out as examples of disruption by Christensen himself, they contest that Christensen’s recipe only fully explains 9 percent them. In the remaining 91 percent, other processes seem to explain technological change much better. Shih’s (2016) account of Kodak’s demise – a widely touted example of disruptive innovation – shows in fine detail how that demise resulted from a much more complex process than what the theory of disruptive innovation lets on. These critiques, I think, are true and should give us pause in thinking about disruption and, at the very least, have us dial down the hype around it. We should recognize it for what it most likely is: one recipe for creative destruction, but one of many possible.

To me, the short-hand way to understand this particular recipe is as a technological change *combined* with the application of the technology in an unattractive market. The combination of these two elements does not imply, as we sometimes think, that incumbent firms are obnoxiously blindsided by change, even though that story may nicely fit a ‘fat and happy’ or ‘bloated bureacracy’ storyline of corporate demise. Contrary to the conventional stories, in-depth histories of incumbents being ‘disrupted’ sometimes show that incumbents *do* know about the emerging technologies, they *do* in some cases understand them and are able to replicate them, they *do* to some extent understand how markets are developing. Incumbents may *underestimate* these developments (which we will return to in subsequent chapters), but the problem also often seems to be that incumbents struggle to *justify* engaging with disruptive innovations. Christensen suggests as much in the second chapter of his book: “although entrants led in *commercializing* disruptive technologies, their development was often the work of engineers at established firms, using bootlegged resources” (p. 43; italics in original; we return to the issue of ‘bootlegging’ in chapter 6.3), but “managers [in the particular case that Christensen examines] made *an explicit decision* not to pursue the disruptive technology. In other cases, managers did approve resources for pursuing a disruptive product -but, in the day-to-day decisions about how time and money would actually be allocated, engineers and marketers, acting in the best interests of the company, consciously and unconsciously starved the disruptive project of resources necessary for a timely launch” (p. 44, italics added).

As Christensen makes clear, engaging with disruptive innovations is often in direct contradiction with what are widely considered ‘best’ business practices and can be difficult because incumbent firms typically depend on their investors’ good graces to operate. Violating best practice can jeopardize investor support, which can de-stabilize firms in various ways. Add to that that investors often prioritize consistent, short-term return on assets rather than what often appears to be high-risk, long-term, explorative projects, especially when investing in incumbents that are meant to be relatively more secure than early-stage ventures. When exploration and exploitation compete for scarce resources and allocation decisions must be justified, exploitation often crowds out especially the very explorative activities that are needed for dealing with disruptive innovation. Rather than have firms undertake this kind of diversification, many investors prefer to manage that portfolio risk themselves. The story of the demise of Blockbuster is especially instructive in this regard (Keating, 2012).

That said, why is it necessary to know about disruption if the theory does not hold water? The straight-forward answer to that is that if you intend to be knowledgeable and conversant about innovation, you need to be familiar with the idea. Knowing the concept, knowing that it is heavily ideologically infused, knowing what it *actually* means and knowing that it does not hold water are all going to be helpful in and of themselves. To talk meaningfully about a topic, we do not just need to know what is true about it, but also sometimes what is *not* true.

The more operational answer is that Christensen's recipe and its reception still does teach us something.

For entrepreneurs, there is a way to think about disruption as a playbook for market entry. Zuckerman (2106) lays out the conceptual argument for this, while Stern & Gans (2018) provide a concise and very practical way to think about the disruption play as one of several entrepreneurial strategies. At the heart of these interpretations is the idea of entering at the low end of a market and scaling up from there, actively speculating in the dynamics of technological improvement (it may be Fosterian improvement or architectural change, as we will return to in 3.2.) and in the myriad different difficulties that incumbents may face in moving to unattractive markets.

For incumbents, it should be read as a warning, not a prediction. Disruptive innovation is something that can happen and that firms should be aware and possibly even anxious about. In this capacity, disruptive innovation runs into the interesting facts of social science that our object study not only changes over time, but actively gets shaped by our theories. Social science theories are not cameras that neutrally depict the world, but engines that contribute to producing and changing that world (MacKenzie, 2006).

Consider Google. In many ways, this is a company of the admired sort that might be ripe for being disrupted. But arguments that the company does several things to actively avoid that, to avoid competitors entering in even low-end or niche markets, are increasingly emerging in popular media. One thing the company does, it appears, is to actively blocks competitors from appearing in Google searches, effectively rendering them invisible to consumers. Second, it seems to actively copy disruptive innovations and give these copies preferential treatment in Google searches, making them highly visible. These two tactics together make market entry by potential disruptors very difficult (Duhigg, 2018). Third, it actively buys up would- be competitors early on to the extent that 'getting bought by Google' is replacing 'disrupting Google' as the preferred end-game plan for many entrepreneurs in the industry. To see why this might be the case, look up the list of mergers and acquisitions by Google's parent company Alphabet. Fourth, and somewhat in a different league, there are suggestions that Google works to actively suppress studies of these business practices (Stoller, 2017). The only good thing to say about that is that Google at least has had the common decency to stop using *Don't be evil* as a corporate motto.

These issues all speak to a larger question of antitrust and government intervention in curtailing the formation of monopolies. For now, note that they are exemplary of how a social science theory can shape what it purports to study. When a theory like Christensen's becomes widely popularized and gets interpreted as a warning, chances are that the warning may be heeded and that incumbents start actively working to prevent disruption, thus making the theory *less* true over time – irrespective of how true it actually was in the first place (e.g. Alvesson & Kärreman, 2007). When, as is the case with innovation, we are dealing with a phenomenon where today's inefficiencies can be tomorrow's opportunities, this is doubly the case, as we will return to.

3.2. Henderson & Clark: Product architectures and organizational architectures

Foster alerted us to how technological discontinuities can devastate incumbents in an industry. With Christensen, we saw how 'disruption' can come from the combination of technological discontinuities and low-end markets. This kind of innovation has a certain dramatic flair in many ways. Discontinuities grab headlines because their novelty is obvious. But other, less dramatized forms of innovative change can – through a different set of mechanisms – bring about equivalent change in industries. In Henderson & Clark's paper, we are going to look at another way for innovation to play out that is much more subtle than what we have previously seen: 'architectural innovations'. An architectural innovation is a "reconfiguration of an established system to link together existing components in a new way" (p. 11). Architectural innovation leaves "the core design concepts (and thus the basic knowledge underlying the components) untouched" (p. 10). Harkening back to Vincenti's treatment of landing gear designs, we are operating at the level of components and subcomponents here, but not looking at change *in* the components. We are looking at how those components are put together, at the product architecture. It is emphatically not about new systems, new components or new S-curves. It is altogether more mundane, focusing merely on the *reconfiguration* of systems and components that already exist. As it turns out, such reconfigurations can bring about the same kind of industrial change as more bluntly novel solutions. Why that is, is going to tell us a lot about organizational dynamics around innovation.

It can be a little difficult to wrap your head around what precisely an architectural innovation is, and there is a good reason for that: it is something that we are unlikely to experience directly as consumers of technology. Henderson & Clark's example comes from semiconductor photolithography, as was the case with the paper by Adner & Kapoor that we read in the previous chapter. Their argument is that the machines that are used to make semiconductors have undergone substantial changes, some of which have been architectural. As a consumer, you will never feel a change like that and if you did, you would never do it with the kind of intensity that industry insiders do. That contributes to the relative obscurity of what is studied here. Henderson & Clark mention some more popular examples, though, like portable transistor radios emerging in the 1950s. A more recent example of an architectural innovation that many may be familiar with is the GoPro camera. The main innovation in that case is not the components (all of those were well understood before the GoPro and were not developed further in any dramatic way). It is the way they are put together to produce a novel functionality.

For all its seeming obscurity, there are two reasons to spend time thinking about this concept. For one, it adds an important nuance to our thinking about different kinds of innovation. Second, it challenges incumbents in a way that is somewhat profound, perhaps even puzzling.

Thus far, we have talked about different types of innovation. There can be innovations in products or production processes, in marketing and organizational practices, at the level of components or systems. Innovations can also be incremental or radical, perhaps even so Schumpeterian that they strike at the center of profits. These distinctions, as also previously covered, are important because innovations of different types operate in different ways – they are part, if you will, of different 'recipes' for different kinds of change in organizations, industries and societies. Henderson & Clark add two concepts to our repertoire of distinctions, with one being much more central to them than the other, and they do this by moving from a one-dimensional to two-dimensional distinction of innovation novelty. In a one-dimensional distinction, an innovation can lie on spectrum from incremental to radical. However, if you think about novelty in two dimensions – whether *core concepts* change and whether *linkages* change – what you get are four types of innovation: incremental and radical innovations that we have already dealt with, modular innovation (which is not the topic of this paper, but important in other ones) and architectural innovation (which is what is central here, as the title gives away). This is what is summarized in the paper's figure 1 and well-illustrated with the example of fans (pp. 12-13). Core concepts describe the components or sub-systems that are part an innovation. Linkages are how those core concepts are linked together to form a system.

Why does this distinction matter, then? What separates an architectural innovation from incremental and radical one?

To answer that question, we need to detour a little bit into what happens to firms once an industry settles on a dominant design. Before settling on a dominant design, Henderson & Clark argue, there will be multiple competing designs that each consist of a set of subsystems and components, and a particular way for those components to be put together – a particular product architecture. Once a dominant design emerges, the architecture of products become largely taken for granted and attention shifts towards improvement in components and subsystems. With the dominant design for e.g. smartphones largely settled, firms now compete to make better smartphones with better screens, faster processors, better cameras, etc. while largely taking for granted how those components will interact. This has an effect on how firms innovate, Henderson & Clark argue, taking their cues especially from ideas around 'organizational routines' (most notably Nelson & Winter, 1982). With a dominant design settled, firms stop exploring alternative designs and begin exploiting the dominant one. This means that their knowledge about alternative designs withers away[7] and that the dominant design begins to shape organizational routines.

The organization comes, in subtle ways, to mirror its product through its development of routine ways of working. The organization moves from fluidity to specificity, not in its production apparatus (as Abernathy would imagine it), but in its innovation processes. Communication channels in the organization come to mirror how the product is designed and the kind of interfaces that the product architecture introduces, both formally and informally. With increasing specificity, development work tends to get organized around the particular components that people work on (e.g. a car's engine, suspension, steering, etc.) and people from different departments tend to interact around the problems that occur where their sub-systems interact (e.g. windshield developers might talk more to window designers than to engine designers). Communication filters emerge, to economize on the amount of information that people need to deal with. If you work on engines, maybe you do not need to spend time hearing about the latest developments in paint application tech-

[7] We will return to the idea of knowledge withering away when we cover Cohen & Levinthal's concept of absorptive capacity.

niques. Efficient strategies for solving problems start to develop, because a product architecture tends to create problems of a particular sort that can be dealt with using solutions of a particular sort. Just like how a dominant design creates specific production systems in the interest of increased efficiency, so does it create specific, routinized and efficient innovation development systems.

As it turns out, architectural innovation creates several problems for this specific, routinized and efficient innovation system. Henderson & Clark describe ‘the Kasper Saga’, noting how an incumbent (Kasper) failed to recognize that a competing product (offered by Canon) actually differed from their products because they looked only at components and failed to notice how a subtle change in the relationship between components allowed the competing product to outperform their own. In this way, the incumbent underestimated the novelty of a solution based on new architectural knowledge, considering the competing product to essentially be a copy of their own – “Kasper conceived [the new product] as a modified [old product]. Like the incremental improvements to the [old product] before it, design of the [new product] was managed as a routine extension to the product line” (p. 25). Relatedly, they also underestimate the difficulties of delivering new products of their own to compete with those of their competitors, misunderstanding the reasons for why their own product could not perform at the level of competitors. As a result, they focused on the ‘wrong’ things as they sought to improve product performance.

What is interesting to note here is that this underestimation of new entrants and new competing products does not, as was the case with disruptive innovations for Christensen, come from an underestimation of how technological improvement might make unattractive markets attractive. It also does not come from underestimating the rate of technological improvement, as would be the case for Foster, or the relative development potential of old and new technologies, as would be the case for Adner & Kapoor. Rather, here we are dealing with what seems to be an actual failure, an actual organizationally induced blindness of sorts.

There are at least two ways to think about that. One is the knowledge that would be necessary to appreciate the novelty and implications of an architectural innovation *simply does not exist* and is very hard to build up in an organization. By implication, it would also be costly to develop and so firms might tend to prioritize it less when not engulfed in an era of ferment. Instead, they might prioritize more exploitable forms of learning. That would be the Marchian interpretation and it would imply that firms would need to continuously explore alternative designs, even if this is potentially inefficient and costly. An alternative interpretation, which is what Henderson & Clark seem to suggest, is that the kind of knowledge that would be necessary to identify an architectural innovation is beyond any one person and thus depends on cooperation between individuals with different knowledge bases. The problem is that this cooperation is impeded by formal and informal communication channels, information filters and problem solving approaches. The implication would be that the firm would need to work to keep these routines at bay and to remain fluid. This would certainly come at the price of lowered efficiency and slower development times.

In both cases, the need to be efficient in the short term is what forces an organization to accept, respectively, a certain level of forgetting and a certain level of inertia. This, in turn, renders it potentially vulnerable to competitors and new entrants. To stay on top of the heap, organizations (mostly implicitly) accept the risk getting caught blindsided and being relegated to an altogether less attractive heap, namely the ash heap of history.

3.3. David: The far reach of general purpose technologies

So far, we have looked at technologies and innovations that shape and, in some cases, transform their industries. Many of them can be important. But some technologies have much deeper implications: they change *many* industries, bring about industrial (sometimes social) revolution, even change the human condition. These are what we call general purpose technologies, or foundational technologies. They are watershed innovations, like the wheel, steam engines, electricity, the concepts of mass production and factories, computers or the internet. Today, technologies like blockchain and artificial intelligence are poised to become tomorrow's general purpose technologies. Just by looking at these examples, it is not hard to imagine that these innovations *must* be conceptually different from those that we have previously met in terms of their characteristics, their effects and especially how those effects are brought about. Those are what David's paper illuminate.

The background for his analysis is what is called the 'productivity paradox'. This term was quite prevalent in economics in especially the 1970s and 1980s and it is, just to be clear, not the same as the productivity dilemma that we talk about in management studies. Rather, the productivity paradox refers what was a genuine puzzle at the time. On the one hand, computers were being widely adopted and were beginning to change people's everyday work in quite dramatic ways. On the other, national statistics were showing a *slowdown* in productivity improvement. In aggregate, everyone appeared to be getting *less, not more, efficient* by working in new ways and investing sometimes heavily in new technology. Robert Solow (who we encountered in passing in chapter one) famously captured the sentiment of the paradox when he stated that "what everyone feels to have been a technological revolution, a drastic change in our productive lives, has been accompanied everywhere, including Japan, by a slowing down of productivity growth, not by a step up. You see the computer age everywhere but in the productivity statistics" (Solow, 1987). As I mentioned in the first chapter, we usually associate innovation with precisely productivity improvements, so you can see why this would be puzzling, even alarming. In the least gracious and curmudgeonly interpretation, maybe computers were just a (really expensive) fashion.

We now know that they were not, but we do also have the advantage of hindsight that David did not.

In David's 1990 argument, he posits that computers belong to a special class of innovations – they are a *general purpose technology* – and therein lies an explanation for the puzzle. To make that clear, he puts the (lack of) effects of computers into historical context by comparing them to the effects of the dynamo. The dynamo and the computer, it turns out, follow broadly similar patterns or, as David puts it "… each of the principal empirical phenomena that make up modern perceptions of a productivity paradox had its striking historical precedent in the conditions that obtained a little less than a century ago in the industrialized West… In 1900, contemporary observers well might have remarked that the electric dynamos were to be seen "everywhere but in the productivity statistics!" " (p. 356).

When electricity was first introduced, it was overwhelmingly used for lighting. Edison invented the electric lamp in 1879, but by the turn of the century only 3 percent of households were using electric lights and – more importantly for productivity concerns – only 5 percent of the power used in manufacturing came from electric sources. The real bottleneck in increasing productivity was the slow pace of adoption of electricity in factories[8]. There are some basic economics of when one might replace a well-functioning machine with a newer one that explain some of this, but the really profound part of the explanation is that electricity did not just mean that a machine shifted from one power source to another. With electricity, the whole idea of what a factory was could change. Factory buildings could change, production lines could be designed in radically different ways, work on the line could change and in turn management practices had to change. "Although all this was clear enough in principle, the relevant point is that its implementation on a wide scale required *working out the details in the context of many kinds of new industrial facilities, in many different locales*" (p. 358, italics added). This required a learning process that was distributed across geographies and people and was highly dependent on costly experiments being undertaken to apply this new and still uncertain technology. That, David posits, explains why general purpose technologies take time and that is why computers took a long time to have discernible impact – they were general purpose by nature and depended on follow-on innovation to achieve their real impact. What mattered was not replacing a typewriter with a computer. What mattered was using a computer to do things that were impossible to do before. Figuring that out required forms of experimentation and learning that were not obvious at the outset.

To put this into terms that are familiar, general purpose technologies exhibit strong characteristics of exploratory innovations and tend to follow an exaggerated S-curve. While the potentials of a technology may be clear[9] (at least in principle), there will often be deep uncertainty about whether those potentials can be practically realized and how that may happen. It will also often depend on the (uncoordinated) choices of many distributed actors. As a result, there will be considerably distance in time between first investments and far-reaching benefits materializing. That means that we will see an S-curve that is drawn out across the x-axis. However, the performance improvements that can accrue directly from a general purpose technology will also tend to exceed those of more conventional technologies dramatically. Relying on these performance improvements are what you will sometimes hear talked about as 'riding the S-curve'. This tends to refer to companies leveraging improvements in technologies without directly driving that improvement themselves. The S-curves that get ridden are often those of general purpose technologies.

Much more dramatically than conventional technologies, general purpose technologies have indirect benefits as well. These primarily come in the form of what are sometimes called 'innovation complimentarities' (Jovanovic & Rousseau, 2005), which describes how a technology enables follow-on innovation. It is generally the case of technology that they build on one another, but a defining characteristic of a general purpose technology is the dramatic extent of follow-on innovation. Consider, for example, the staggering amount of innovation that has been built on top of the internet. While important, the original use of the internet for connecting respectively military and academic networks together is largely eclipsed by the way innovations in how internet technology have transformed virtually all aspects of life and commerce in greater and greater parts of the world, and will probably continue to do so. This is precisely because of the opportunities of innovation complementarities. A more current example of this is the (at the time of writing) highly hyped Blockchain technology which is also poised to enable follow-on innovation in a just astounding range of areas (Iansiti & Lakhani, 2017).

---

[8] There is a wonderful talk by Jeff Bezos on how electricity transitioned from being a 'ligthing-only' technology to the way we power everything, and what that teaches us about the web's future impacts. A link is provided under 'Recommended videos and podcasts'.

[9] As we will cover in the subsequent sub-chapter describing Shane's ideas of entrepreneurial opportunity discovery, they often are not particularly clear.

Through their direct and indirect benefits, general purpose technologies often get described as important 'engines of economic growth', driving whole eras of economic progress (Bresnahan & Trajtenberg, 1995), albeit with considerable lag between their invention and impact. In this sense, all the ways we make use of general purpose technologies end up shaping our lives. But the impacts of general purpose technologies actually extend even further, shaping even the human condition, as I somewhat prosaically claimed at the opening of this sub-chapter. As Rosenberg & Trajtenberg show in their study of the Corliss Steam Engine, this particular general purpose technology not only transformed how production happened, but also *where* it happened (2004). Before the steam engine, factories had been dependent on water for power and thus factories tended to be located outside of urban centers, close to large rivers. With the advent of the steam engine, this constraint was lifted and it became possible to locate factories closer to consumer markets and closer to thick labor markets. Suddenly, cities could be centers not only of trade but also of large-scale production. Industrial work changed from being something that farm girls did before getting married to being something that women and men moved to towns to do for all of their working lives. The impact of that change is hard to overstate: increased urbanization is one of the defining points of the industrial revolution and it played a massive role in the emergence of new political ideologies and social classes that still matter today. Later, this urbanization would play a central role in many other changes that shape how we live today.

3.4. Shane: Opportunity discovery and making technologies general

All the forms of innovation that we have covered in this chapter depend on opportunity recognition, on someone recognizing that a new solution can somehow create value. This is clearly the case for general-purpose technologies. When a general purpose technology emerges, the full extent of its applicability is not known. The 'killer app' for a particular technology is often not the one that the inventor envisages. Rather, the full range of applications that turn a technology into a general one have to discovered and figured out. It is also the case, but perhaps less clearly, with disruptive and architectural innovations. A disruptive innovation depends on someone recognizing that a particular early-stage technology can be used to create value for a particular group of 'low end' customers who are currently overserved by established solutions and could be satisfied with a poorer performing one. An architectural innovation depends on someone recognizing that additional value – whether in the form of new functionality or increased performance – can be created by reorganizing a set of known components in a novel way. The question that motivates Shane's paper is *how* those opportunities get recognized.

His study pits three competing views against each other. The baseline assumption of mainstream economic thinking ('neoclassical equilibrium theories') is that markets are composed of maximizing agents and that markets function well. In this world, there is little opportunity for entrepreneurship: "no one can discover a misalignment that would generate an entrepreneurial profit because, at any point in time, all opportunities have been recognized and all transactions perfectly coordinated" (p. 449)[10]. What drives entrepreneurship is instead personal preference, typically those with a greater preference for risk or uncertainty or, perhaps less conventionally, greater disutility from subordination to managers. The baseline assumption in the psychological position on entrepreneurship is not unrelated. The view here is that there are stable human traits and that

[10] This – as we have just explored in the context of general purpose technologies – is *empathically* not a realistic assumption to make. Economists tend to have more nuanced views than Shane gives them credit for, even if those views do not always filter into the way their theoretical models work. It is more fair to say that mainstream economists assume that opportunities exist and can be readily be discovered through investment in 'search'. They may exist on 'rugged landscapes' (Levinthal, 1997), but investment will uncover them.

these are what determine the decision to enter into entrepreneurship[11]. It is not so much a question of preference, but rather a question of "enduring human attributes" (p. 449) or "superior information processing ability, search techniques, or scanning behavior" (p. 449). The baseline assumption of Austrian economic theories is that "markets are composed of people who possess *different* information (Hayek, 1945). The possession of idiosyncratic information allows people to see particular opportunities that others cannot see, even if they are not actively searching for such opportunities" (p. 449)[12].

Shane's argument is that the third view, Austrian theories, provides the most compelling answer. His argument is that opportunities get discovered largely through *serendipity* (i.e. without much deliberate search) by people who possess particular prior knowledge *relevant to that particular opportunity* (i.e. knowledge makes only *certain* opportunities obvious) and that people's decision to exploit a particular opportunity depends *on the opportunity* (i.e. it is not a feature of people's inherent mental attributes). The knowledge needed to recognize an opportunity, in other words, *is* highly distributed. That has considerable implications for how we think about innovation management as an intra-organizational phenomenon that we will be returning to at length in subsequent chapters. Leaving that for the future, we will now examine *how* Shane arrives at this conclusion.

Because he starts out with a clear intention to demonstrate that Austrian theories do a better job of explaining opportunity recognition than mainstream economics and psychology, Shane initially draws out a set of propositions based on an Austrian perspective and in disagreement with alternative ones. These are: 1) All individuals are not equally likely to recognize a given entrepreneurial opportunity; 2) People can and will discover entrepreneurial opportunities without actively searching for them; 3a) People's prior knowledge about markets will influence their discovery of which markets to enter to exploit a new technology; 3b) People's prior knowledge about how to serve markets will influence their discovery of how to use a new technology to serve a market; 3c) People's prior knowledge of customer problems will influence their discovery of products and services to exploit a new technology (pp. 451-452, graphically summarized on p. 453). It may spoil the plot somewhat, but those propositions all turn out to be supported by the analysis.

To test these propositions, he examines the entrepreneurial opportunities identified for 3D-printing which is at my time of writing (2018), depending on your temperament, possibly a general purpose technology right about the point of having 'all hell break loose' (Foster, 1986, p. 31). Basically, he looks at every entrepreneur that contacted the owner of the patent on this technology with an interest in commercializing it. We know that general purpose technologies can be a long time underway, but it is interesting to note that in the ten years after the technology was developed, only eight companies sought to apply it to commercial ends. The opportunities range from printing conceptual models for architects to artificial bones for surgery to ceramic molds for industrial metal casting.

[11] The view that human traits are stable is, to be clear, not a position that is shared across *all* of psychology. There seems to be agreement that *some* traits are relatively time-stable, but also that many are not. Some will also argue that traits are highly context-dependent and subject to change given appropriate 'stimulus', rejecting much of the assumption of stability.

[12] A range of other assumptions are embedded in Austrian economic perspectives aside from distributed information and some of these I will return to in Appendix I.

Interestingly (and in accordance with Shane's propositions), the individuals who did pursue entrepreneurial opportunities around the technology were highly limited in their opportunity recognition. They recognized only the particular opportunity that they pursued and did so largely because they possessed some relatively idiosyncratic experience, typically from work experience in the market that they intended to address with their particular solution (i.e. their former professional selves are their current customers). Because of this knowledge, they 'just' recognized the opportunity. The alternative way that you could think about how technology gets commercialized through systematic search largely does not happen: no one started the technology and investigated different applications by doing market research of various kinds They largely know the markets that they will address in advance and that is what makes an opportunity appear and appear attractive. And of course – this being the death knell to the mainstream economic argument – no one can possess enough knowledge and enough experience to identify all possible, or even relevant, 'Need-solution pairs' (Von Hippel & Von Krogh, 2015).

In itself, this is interesting mostly in a quizzical sense, as a bit of trivia from the history of technology. The wider implications of the argument, however, are what is truly interesting. They concern the way that innovations are evaluated and the way that different people are going to see different potential and value in different innovations, because those evaluations are going to shape what decisions get made.

If we think about new potential new entrants to an industry, it makes it plausible to think of innovation as eminently *likely*. There are, in Shane's Austrian conception of the nature of entrepreneurial opportunity, vast opportunities. You will sometimes hear the advice, originating I believe with Peter Thiel (2014) and now pervasive in tech jargon, that you should look to build 'the valuable company that nobody else is building'. In a world of perfect information, it is per definition hard to imagine that you could realistically be the first person to identify an opportunity and a solution to it – if that company does not exist, it would probably be because there is no way to build a profitable business around it. But in an Austrian world, your insight might actually just be that unique. You may be the first person to have the combination of prior knowledge, insight in a technology and an idea for exploiting it. The same prior knowledge that affords an advantage in recognizing one particular opportunity, however, can also easily create a disadvantage in recognizing *alternative* opportunities, because it creates a 'knowledge corridor' (p. 452) that both enables and constrains opportunity recognition.

If we think about Shane's perspective in a more organizational sense, the effects of this 'corridor' become more one-sidedly constraining. Imagine that an idea originates somewhere, anywhere in a large organization. The person who originates it sees tremendous value in it, possibly because it solves a problem that she has herself experienced or because she works with a new technology that others have yet to fully understand. For that innovation to have far-reaching impact, it will have to be shared with many others who will have to see similar value in it (the same thing would be the case for our entrepreneurial start-up above, who will often have to find venture capital funding for the idea). Now, almost per Austrian definition, others will *not* see that value because the knowledge required to appreciate it is *not* going to available to them. In an organizational setting where cooperation and coordination and resource allocation requires many more people to have a shared evaluation, in an Austrian view innovation may well be very likely but recognition of that innovation by others highly unlikely, leaving many ideas underdeveloped relative to their potential. We will return to this conundrum in the subsequent two chapters as we look at opportunities for organizations to 'work around' the Productivity Dilemma and at the challenges of managerial control over innovation processes.

Chapter IV:
Creativity

4.1. Ward: Structured imagination

4.2. Sternberg: The nature of creativity

4.3. Czikszentmihalyi: A systems view

Chapter 5:
Organizing around the Productivity Dilemma

5.1. Adler et al: Re-thinking the trade-off between exploration and exploitation

5.2. Brown & Duguid: Organizational learning and inevitable innovation

5.3. Cohen & Levinthal: Absorptive capacity and the effects of R&D

Chapter VI:
Managerial control in innovation processes

6.1. Cooper: Innovation through stages

6.2. Mouritsen et al: Management accounting and economizing innovation projects

6.3. Augsdorfer: Bootlegging and reactive organizations

6.4. Rennstam: Control by other means

Chapter VII:
Bringing new technologies to market

7.1. Lieberman & Montgomery: First-mover advantages and risks

7.2. Navis & Glynn: The emergence of market categories

7.3. Moore: Beachhead markets and The Chasm

Chapter VIII: Innovation's consequences

Texts in this chapter:

Autor, D. H. 2015. **Why are there still so many jobs? The history and future of workplace automation.** Journal of economic perspectives. 29(3): 3-30.

Bessen, J. 2014. **Learning by doing: The real connection between innovation, wages, and wealth.** (Ch. 5-6). Yale University Press.

Lu, S. F., Rui, H. & Seidman, A. 2017. **Does technology substitute for nurses? Staffing decisions in nursing homes**. Management Science. 64(4): 1842-1859.

Moretti, E. 2012. **The new geography of jobs.** Houghton Mifflin Harcourt.

8.1. Autor: Substitution and complimentarity

8.2. Bessen: Technology, wages and Engel's pause

There are lots of dilemmas in capitalism, the fact of creative destruction being an obvious one. Another and closely related one has to do with the gains from technology. Who benefits from technological change? Recalling McCloskey's chapter (discussed in this book's section 1.2.), there is no doubt that all the innovation that capitalism contributes to bringing about *dramatically* increases human living standards. We are incredibly wealthier today than we were two hundred years ago. And yet, when we look for instance at living standards during the Industrial Revolution – that great leap forward in human history – they were dismal. When Victor Hugo and Charles Dickens wrote their most famous works, they were describing very real poverty, destitution and exploitation. When Karl Marx and Friedrich Engels wrote about the workings of capital and the conditions of the working classes and co-authored the Communist Manifesto, their analyses were mostly right. Workers were producing more and more, but gaining almost none of the fruits of their more and more machine-complemented and efficient labor. We see similar patterns today, as productivity grows more rapidly than wages, the middle classes are shrinking and an increasing share of wealth seems to be concentrating with "the 1%". How do we make sense of this dilemma?

There are two questions in play here. Put somewhat simply, they are about the size of the 'economic pie' and about how that pie is divided. It is quite clear that capital is complementing labor, increasing efficiency and allowing new jobs to be invented, as Autor argues. In the process, the economic pie grows because we become able to produce more with the same resources. It is less clear that all of us benefit from this growth. At least in the current historical moment, the benefits of growth seem to be concentrated in particular parts of the income distribution, in particular demographics and in particular geographies. As Schumpeter argued, we may on average be made better of *as consumers* of good as services, but in the simultaneous capacity as workers *producing* goods and services experience increased variance in outcomes, as Marx would have expected.

To understand the dynamics of why this happens, Bessen's book is an analysis of the wage developments of textile workers during the Industrial Revolution. To be more precise, he looks at textile workers in Lowell (Massachusetts, USA) and their wages from 1830 to 1890. This may seem offbeat, but at the time Lowell was quite the Silicon Valley. At the time, the textile industry was as powerful a force as 'tech' is today and Lowell was at the bleeding edge of the technological dynamism that drove this industry. Also, between 1830 and 1860, textile workers in Lowell experienced a situation that is strikingly similar to what we see today. They saw their productivity rise incredibly, but saw little increase in their wages. From 1860 onwards, however, their wages began to rise. In this sense, they follow the pattern of what Robert Allen (2009) has called "Engel's Pause" – an extended period of productivity gain without wage increases. Looking in detail at this particular site allows Bessen to observe micro-level patterns that would not be visible at more aggregate levels. He can see, for instance, how productive individual workers were, what they were paid, how much companies invested in new technology and all kinds of other details. That allows him to offer an explanation of *why* Engel's Pause occurs and *what* it is that brings it to an end.

To understand why wages were stagnant at the beginning of the period, we have to understand what defined technology in the industry at its early stages, during what Anderson & Tushman would call its era of ferment. You have to imagine that the textile industry at this point is relatively new and small, and geographically scattered (Rosenberg & Trajtenberg, 2004). Workers would typically be farm girls, moving from their country homes to a particular factory where they would work and board, much like migrant workers in contemporary Chinese factories. More importantly, that is no dominant design for production machines, as you would expect from an era of ferment. As a result, machines were idiosyncratic to the particular factory. As workers gained experience in a particular factory, they would often become much more productive – if they did not become more productive, they would tend to leave the factory and return to the farms they came from. However, the experience that they gained was *only* valuable in *that* particular factory, because another factory would use different machines. In economics jargon, the skills they learned on the job were 'firm-specific' (pretty much worthless in another factory) and the labor markets they operated in were 'thin' (with few employers). Taken together, this gave factory workers poor alternative employment opportunities. And that weakened their ability to negotiate better wages, irrespective of their productivity. The market for textile labor was a buyer's market.

That started to change around 1860 as the industry reached increasing maturity. The industry grew dramatically to elastic demand (Autor, 2015) and this lead to more employers. It also grew increasingly geographically concentrated. Both of these things, you would think, would make the labor market 'thicker', increase workers alternative employment opportunities, strengthen their bargaining position and ultimately lead to higher wages. The real driver of change, however, was that technological change in the industry *slowed down*. As you would expect following Andersen & Tushman, an era of ferment closes with the emergence of a dominant design. With that emergence comes not only reduced competition between designs, and therefore less intense innovation, but also increased standardization. When machine designs stopped changing rapidly, it made increasing sense for workers to invest more in learning how to use a particular machine, e.g. by going to school and learning it or by acquiring more advanced skills on the job. Also, it turned textile work into something that you *could* actually learn in professional schools, because it could actually be taught as a standardized subject. When more factories started using similar machines, skills learned either in school or at a particular factory ceased to be firm-specific and became transferable. That made skills more valuable and made investments in building skills much attractive. When it started to make sense to invest in acquiring skills, it also began to make sense to do a job for longer. In Lowell, that meant that textile work stopped being done by farm girls, who would typically work for only a few years. Instead, it was something that city dwellers did for a livelihood and thus deeper investments in developing skills that you could use for ten or twenty years became sensible, making workers more productive. It also meant that workers remained in the cities that had many potential employers, making it more lucrative for firms to set up shop there. All in all, this allowed workers to gain a stronger bargaining position and demand higher wages, allowing them to capture a larger share of value created by their increased productivity. The slowing down of technological change in the textile industry made the market for textile labor more of a seller's market.

While interesting in itself, it is an interesting aside to this description that while Engels' work on the conditions of the working class was very true at the time it was published in German (1845), it was remarkably *less* true when it was published in English (1885). Put another way, both Marx and Engels were right – until they were wrong. Like many other theories, they became wrong because the world changed and may be right again.

If we extrapolate the pattern Bessen describes to today, living as we do amidst dramatic technological change on multiple frontiers, what is in store for us? Bessen's argument would clearly be that we are living in a new Engel's Pause and that we should expect a period where productivity increases much more than wages. We can also reasonably expect this period to be long as we wait for the gale of creative destruction to calm and for the myriad new technologies that surround us to stabilize and standardize. Especially for workers in the lower part of the wage distribution, things may very well get worse for a long time before they get better. They will, however, get better eventually and when that happens, not only will we (if Bessen is right) see that living standards have increased *on average*, but also that the variance in outcomes is reduced and the gains from technology become widely shared. Until that happens, we can hope that the social unrest that inevitably follows increased variance in social outcomes will not spawn political agendas that seek to stop technological change, but instead seek to limit its negative impacts on those most adversely affected. As McCloskey made clear, ideologies can change and so can the ideology of innovation.

8.3. Lu et al: Substitution as a choice

8.4. Moretti: Winner-take-all Geographies

Chapter IX:
The future of innovation management

9.1. Altman et al: Innovating in the age of the internet

9.2. Lifsitz-Assaf: From problem-solvers to solution seekers

9.3. Jeppesen & Lakhani: Broadcast search

9.4. Arora et al: Markets for technology

Chapter X:
Innovation without organizations

9.1. Von Hippel: The consumer innovator

9.2. Benkler: Innovation in the commons

9.3. Davis: The end of the Corporation?

Chapter XI:
Concluding thoughts

References and futher reading

Abernathy, W. J. 1978. **The productivity dilemma: Roadblock to innovation in the automobile industry.** John Hopkins University Press.

Adner, R. & Kapoor, R. 2016. Innovation eco-systems and the pace of substitution: Re-examining technology S-curves. **Strategic Management Journal**. 37(4), 625-648.

Allen, R. C. 2009. Engels' Pause: Technical change, capital accumulation and inequality in the British industrial revolution. **Explorations in economic history.** 46(4): 418-435.

Alvesson, M. & Kärreman, D. 2007. Constructing Mystery: Empirical matters in theory development. **Academy of Management Review.** 32(4), 1265-1281.

Anderson, P. & Tushman, M. L. 1990. Technological discontinuities and dominant designs: A cyclical model of technological change. **Administrative Science Quarterly**. 35(4), 604-633.

Bloom, N., Jones, C. I., Van Reenen, J. & Webb, M. 2018. **Are ideas getting harder to find?** NBER Working paper No. 23782.

Bresnahan, T. F. & Trajtenberg, M. 1995. General Purpose Technologies: 'Engines of growth'? **Journal of Econometrics.** 65(1), 8-108.

Brynjolfsson, E. & McAfee, A. 2014. **The Second Machine Age: Work, progress and prosperity in an age of brilliant technologies.** W. W. Norton & Company.

Burns, T. & Stalker, G. M. 1961. **The management of innovation**. Tavistock.

Christensen, C. 1997. **The Innovator's Dilemma: When new technologies cause great firms to fail.** Harvard Business Review Press.

Colaizzi, P. 1978. Psychological research as a phenomenologist views it. In: Valle, R. & King, S. (eds.) **Existential-Phenomenonological Alternatives for Psychology.** Oxford University Press.

Cusumano, M. A., Mylonadis, Y. & Rosenbloom, R. S. 1992. Strategic Maneuvering and mass market dynamics: The triumph of VHS over BetaMax. **Business History Review.** 66(1), 51-94.

David, P. A. 1985. Clio and the economics of QWERTY. **The American economic review.** 75(2), 332-337.

David, P. A. 1990. The dynamo and the computer: An historical perspective on the modern productivity dilemma. **The American Economic Review.** 80(2), 355-361.

Dosi, G. 1982. Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. **Research Policy.** 11(3), 147-162.

Duhigg, C. 2018. The case against Google. **The New York Times Magazine**. February 20, 2018.

Foster, R. 1986. **Innovation: The attacker's advantage**. Summit Books.

Gordon, R. J. 2016. **The rise and fall of American growth: The U.S. standard of living since the Civil War.** Princeton University Press.

Henderson, R. M. & Clark, K. B. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. **Administrative Science Quarterly.** 35(1), 9-30.

Iansiti, M. & Lakhani, K. R. 2017. The truth about Blockchain. **Harvard Business Review.** 95(1) **(**January-February 2017), 118-127.

Jovanovic, B. & Rousseau, P. L. 2005. General purpose technologies. In: Aghion, P. & Durlauf, S. N. (eds.) **Handbook of economic growth**. Elsevier.

Keating, G. 2012. **Netflixed: The epic battle for America's eyeballs.** Portfolio.

King, A. A. & Baatartogtokh, B. 2015. How useful is the theory of disruptive innovation. **MIT Sloan Management Review.** September 15, 2015.

Lepore, J. 2014. The disruption machine: What the gospel of innovation gets wrong. **The New Yorker.** June 23, 2014.

Levinthal, D. 1997. Adaptation on rugged landscapes. **Management Science.** 43(7), 934-950.

Lewinson, M. 2008. **The box: How the shipping container made the world smaller and the world economy bigger.** Princeton University Press.

MacKenzie, D. 2006. **An engine, not a camera: How financial models shape markets.** MIT Press.

March, J. G. 1991. **Exploration and exploitation in organizational learning.** Organization Science. 2(1), 71-87

McCloskey, D. N. 2010. **Bourgeois Dignity: Why Economics can't explain the modern world**. University of Chicago Press.

McGraw, T. K. 1995. **Creating modern capitalism: how entrepreneurs, companies and countries triumphed in three industrial revolutions.** Harvard University Press.

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

Moore, G. A. 1991. **Crossing the Chasm: Marketing and selling disruptive products to mainstream consumers**. Collins.

Nelson, R. R. & Winter, S. G. 1982. **An evolutionary theory of economic change.** Harvard University Press.

Raffaeli, R. Forthcoming. Technology reemergence: Creating new markets for old technologies, Swiss mechanical watchmaking 1970-2008. **Administrative Science Quarterly.** *Published ahead of print.*

Rosenberg, N. & Birdzell, L. E. 1986. **How the West grew rich: The economic transformation of the industrial world.** I.B. Tauris.

Rosenberg, N. & Trajtenberg, M. 2004. A General-Purpose Technology at Work: The Corliss Steam Engine in the late Nineteenth-century United States. **The journal of economic history.** 64(1), 61-99.

Schumpeter, J. A. 1942 (2008). **Capitalism, Socialism and Democracy** (Chapter 8). Harper Perennial Modern Thought Edition.

Shane, S. 2010. Prior knowledge and the discovery of entrepreneurial opportunities. **Organization Science**. 11(4), 448-469.

Shapiro, C. & Varian, H. 1998. **Information Rules: A strategic guide to the network economy**. Harvard Business Review Press.

Shih, W. 2016. The real lessons from Kodak's decline. **MIT Sloan Management Review.** May 20, 0216.

Solow, R. M. 1987. We’d better watch out. **New York Times Book Review.** July 12, 1987, p. 36.

Stern, S. & Gans, J. 2018. **Entrepreneurial Strategy.** Available for download from www.entrepreneurial-strategy.net.

Stoller, M. 2017. Citizens against monopoly. **Huffington Post.** Posted August 30, 2017; updated August 31, 2017.

Thiel, P. 2014. **From Zero to One: Notes on start-ups, or how to build the future**. Virgin Books.

Vincenti, W. G. 1994. The retractable airplane landing gear and the Northrop “anomaly”: variation, selection and the shaping of technology. **Technology and culture**. 35(1), 1-33.

Von Hippel, E. & Von Krogh, G. Crossroads – identifying viable “Need-solution pairs”: Problem solving without problem formulation. **Organization Science**. 27(1) 207-221.

Weber, M. 1930. (2001). **The protestant ethic and the spirit of capitalism.** Routledge.

Zuckerman, E. W. 2016. Crossing the Chasm to disruptive innovation. **MIT Sloan Management Review.** March 15, 2016.

Recommended videos and podcasts

Jeff Bezos: “The electricity metaphor for the web’s future”. TED Talk, 2003. Link.

Appendix I: Two contrarian views

I.1. Austrian Economics: Emergence, the nature of knowledge and the lure of Scientism

I.2. Critical Theory: Power, ideology and distorted visions

I.3. A synthetic interpretation

Appendix II: Using this book (a guide for teachers, an explanation to students)

The purpose of this book has been to offer a somewhat different take on how innovation management, and lots of other topics in management, tends to be taught. The book is trying to be something different from especially the standard format of textbooks, as I mentioned in the preface. It is, however, also written to enable a somewhat different way of teaching. In much the same way that I have strained relationship with the standard textbook format, I tend to find that the 'standard' way of teaching in business schools misses many opportunities for learning. I also find that a different approach, centered on what I call 'the long class' is not only viable but desirable. Because the book is written specifically to support this approach, this appendix is going to outline the approach and how the book fits into it.

The standard way that teaching happens at the introductory level tends, somewhat simplistically put, to follow with few variations a standard recipe. For each class, students are assigned a set of readings, either textbook chapters or academic papers. Most students read them, but for good reasons few know what to look for and what to focus on. Students will then go to class, where a lecture is held. This can be more or less interactive and may involve elements of group work or case work, but one-way communication by a lecturer using PowerPoint-slides tends to play a significant role in most courses. During this lecturing, the median student looks at his or her laptop screen for a sizable portion of the time, ideally taking notes and less ideally whiling away the time. Those slides that the lecturer used may get distributed to students, who will revisit them as they prepare for exams. This gets repeated for the following class, and thus the semester progresses. Students do receive feedback, but most of it comes when they receive their exam grade and are on their way to next semester's subjects.

It is certainly true that the lecture format can be a genuinely good and highly productive learning experience, especially when practiced by great lecturers. It has, after all, been with us since the invention of the university. It is certainly also true that the monotony of the lecture is easy to shake up with and that this can create tremendously engaging learning experience. But fundamentally, I cannot help thinking that technology *ought* to allow us to teach in different ways. That, after all, is what technology does: it opens up new possibilities for innovation.

Hal Varian gets quoted for saying "If you want to understand the future, look at what rich people do today". If we follow this line of thinking, we should be asking how university teaching happens today in institutions where resources are plentiful. Harvard Business School certainly fits that bill and teaching there does get done in a different way. Instead of lectures, case-based teaching is the *de rigueur* method. In many business schools and management classrooms, you will also see case work being used. However, an alternative approach – and one that technology might be well-suited to democratizing – is what gets done at the universities in Oxford and Cambridge. While lectures do feature in these institutions, the method at the center of their pedagogy is the tremendously resource-intensive *tutorial system*.

The tutorial system is based around students meeting regularly with faculty in groups of maybe three or fewer students. For each meeting, students will read a certain set of texts and prepare some form of written product. In meetings, students engage in discussions with their peers and the faculty member, thus practicing their verbal skills at communicating, applying and critiquing ideas. This affords plentiful feedback. This, I think, is a more interesting model for how technology-enabled higher education might happen.

To approximate this model in a context of somewhat more limited resources, I try to organize my teaching around the concept of 'the long class'. You should see this as a contrast to the 'short class' that is implicit in what I somewhat crudely called the 'standard' model of management teaching. The long class, as I imagine it, has several components and all of them are intended to increase the depth of student engagement with academic texts. Deep engagement is going to be more challenging, and so rather than just 'piling in' with more material, the challenge of making classes 'longer' is to accompany higher levels of challenge with higher levels of support. To this end, these are the steps that I tend to encourage students to follow as they prepare for class:

1) Read the preparation guide for the class.
2) Read one of the assigned texts
3) Read the accompanying chapter in this book
4) Repeat 3) and 4) for the remaining assigned texts, focusing particularly on one of them – you (the student) will be discussing this particular text in class.
5) Listen to a podcasted lecture.
6) Write a 200-word paragraph about something that you (the student) believe to be true about the topic that you have read about.

An exemplary preparation guide and the podcasts that I use can be found on my blog:
www.TheGale.blog/FoundationLectures

Writing 200-word paragraph is an alternative to writing the short essays that are usually part and parcel of a tutorial mode of teaching. It is also a really helpful way to build a practice of academic writing and something that I encourage students to practice. The approach to writing that I tend to endorse gets discusses in a series of interviews with Thomas Basbøll that you can also find on my blog:
www.TheGale.blog/AcademicWriting

In class, the two hours that tend to comprise a scheduled session gets spent on discussions of texts and no time gets spent on lecturing. Each paper will be allotted an equal amount of class time for discussion (ca. 25 minutes) and all students will have read one text with an eye to discussing it in class. Discussions tend to follow a 'fish bowl' format, where the lecturer (me, in my case) and those students who have prepared a particular text sit at a table in the middle of the room and discuss it, with the remaining students observing the interaction. At the table, discussion can be relatively free-flowing and focus on those issues that students find hard or those questions that they find interesting. Depending on class size and attendance, the number of discussing students can vary quite a bit and that of course produces different dynamics. My experience is that students attain a deeper mastery of the ideas presented because they filter it through their minds much more intensely than in conventional courses, because they do not just read a set of texts, but also read interpretations of them, see connections, and both talk and write more than in the 'standard' model of management education. There is, thus, very little 'bullshit' in the course, in Colaizzi's sense (1978).

It is in this context of a ‘long class’, as a variation on what often gets called a ‘flipped classroom’ pedagogy, that this book is meant to be used. I consider it supplementary to readings of original texts and for this reason I do not assign the book as required reading – it is a help to making sense of the sometimes quite difficult texts that I write about and to seeing them in connection to other texts. In this sense, the book can hopefully be a resource also in courses that use only some of the material that I cover or that understands innovation management in a different way than me. That is part of the reason why the book is freely available – to allow it to be used as a complementary resource, rather than as an independent authority.

There is of course a way to learn about innovation management theory that ‘just’ involves reading the book and you would (I hope, at least) learn something about innovation by doing so. But really my intent is not to be offer an exhaustive account that can stand alone or even be fully balanced. In that respect, reading this book only will give you a view of innovation management theory that is somewhat skewed *by design*. Reading it as a complement to the papers that I present, rather than as a substitute for them, should make them more accessible and more open to critical engagement.