Notes and Reflections

PHILOSOPHY OF

TECHNOLOGY

Daniel Little, University of Michigan-Dearborn

Chapter 1

INTRODUCTION

Section 1

TEACHING AND LEARNING PHILOSOPHY

What is it that we expect students to learn when we teach philosophy? Is philosophy an arcane and charmingly useless vestige of a nineteenth-century university education? Or does it have something crucial to add to the liberal education of the twenty-first century -- whether in the arts and sciences or in pre-professional schools?

Philosophers would probably answer this question in a wide variety of ways. In my own case, I have several high-level goals in mind when I approach a new group of undergraduate students in philosophy. I hope to help them to develop in several ways:

- to gain a set of intellectual skills: analysis, reasoning, clarity of thinking and exposition, open-mindedness and a readiness to try to see a problem from multiple points of view
- to learn some of the developed approaches to "philosophical" problems: knowledge, ethical behavior, individual rights, social justice, the authority of the state, the nature of rationality, the meaning of human life
- to gain an engaged involvement in some great thinkers and their theories and reasoning in detail
- to gain some meaningful acquaintance with some important philosophical theories (utilitarianism, empiricism, mind-body materialism, ordinary language philosophy ...)
- to gain an ability to see the connections between philosophical reasoning and real human problems -- scientific knowledge, addressing poverty or racism, resolving conflicts of value or conflicts of interest or desire, ...

Most of these goals have to do with developing intellectual capacity -- imagination, reasoning, analytical ability, critical capacity to probe behind ordinary assumptions

-- more than gaining specific bits of knowledge about the history of philosophy. Students are exposed to pieces of philosophical traditions that result in exam questions such as these: What was Anselm's ontological argument? What was Russell's paradox? What were the higher pleasures according to J. S. Mill? But the real learning goal isn't that the student should have the ability to draft a short Wikipedia entry on one of these topics. Rather, the goal is that he/she has enough of a set of analytical and critical skills so that she can pick up a philosophical problem; explore and develop the problem with insight and imagination; consider a variety of ways of addressing the problem; and put forward a philosophical argument that attempts to resolve the problem. The student needs to learn how to think -- philosophically, imaginatively, and critically.

It is true, of course, that being able to formulate and resolve a philosophical problem requires a degree of acquaintance with the systems and theories that previous generations of philosophers have brought to bear on the problems they raise. So it is important to have grappled seriously with Anselm, Russell, or Mill. And this means taking seriously the positions these philosophers and others advanced and the intellectual frameworks within which they reasoned. So a degree of knowledge of some of the fields and traditions of philosophy is an important intellectual attainment for a philosophy student. But the goal of pursuing this knowledge is not so that the student can become a mini-expert on Anselm or Russell; rather, the goal is to broaden the set of intellectual frameworks and reference points on the basis of which the student's philosophical imagination can address new problems.

This approach addresses one of the large dichotomies that we have to face in designing a university curriculum: the split between "intellectual skills and capacities" and "mastery of content". My position puts primary emphasis on the former over the latter. One might ask, in good philosophical fashion, why we might want to make this choice? My own reason has to do with the highest goal I think universities ought to pursue: to help their students to gain a rich range of skills, tools, and intellectual resources on the basis of which they can address the widest range of problems they will face in their civic and professional lives. When a philosophy student graduates, attends law school or business school, and enters the world of professional activity, he/she may not be able to reproduce specific arguments from the course she took in epistemology or the philosophy of science. But what we hope is that the challenge of working with those arguments as an undergraduate, challenging and dissecting the assumptions the philosopher made, and considering alternative solutions to the problem, will have given him/her a broad intellectual range and acuity, and a flexible and imaginative ability to think through a set of issues. And, we would hope, these skills are highly transportable, from the context of philosophy to the practical intellectual challenges of being a good doctor, lawyer, or engineer. Ultimately the intellectual capacities of imagination, analytical ability, critical insight, and intellectual rigor are the best and most enduring attainments of a good liberal education.

This goal has to do with intellectual capacity and imagination. But we have another and equally important goal as well in designing a university education or a philosophy course. This is the goal of helping our students become engaged and morally motivated members of the organizations and communities to which they belong. We would hope that our students have cultivated an ability to think independently and seriously about the issues of social justice and personal conduct that arise in the society that they are helping to constitute; and we would hope that they have acquired some of the components of personal seriousness that lead them to act with conviction on the basis of their moral ideas. The transition from narcissism to engagement is not an automatic or inevitable one, and a suitable learning environment in the university can have a large impact on this process of personal development. So my hope in my own philosophy classroom is that students will have an opportunity to explore and challenge their own moral ideas; to come to see how the contemporary world measures up with respect to those ideas; and to see that their own engagement in issues of community, justice, and social progress can make a meaningful difference in the state of their world.

Some of this process of critical self-discovery can happen in the classroom. But some of it is best stimulated by the other activities that can help students get engaged in the important social issues of their day -- poverty alleviation, literacy, racial disparities, etc. Involvement in organizations such as Habitat for Humanity or Amnesty International can give students a genuine understanding of the needs their world presents to them, and the difference that their engagement can make. And the teamwork that unavoidably accompanies all these activities gives a concrete illustration to the student of the value of collaboration.

This line of thought converges with one of the common refrains of current thinking about pedagogy: the idea of the student as an "active learner." As Socrates and Habermas illustrate in the images above, a very large part of teaching philosophy is the challenge of getting the student to think for himself/herself. The student needs to take on the intellectual challenge as a serious one; and he/she needs to expend the real mental effort required to understand and deal with the problem. This can't be distilled into an artful lecture by the professor; rather, it seems to require dialogue and intellectual exchange. The student needs to be engaged in the debate; and he or she needs to be brought to see the stakes of the issue.

Chapter 2

WHAT IS TECHNOLOGY?

Section 1

THE DOMAIN OF PHILOSOPHY OF TECHNOLOGY

Is there such a thing as “philosophy of technology”? For that matter, is there a “philosophy of cooking” or a “philosophy of architecture”? All of these are practical activities – praxis – with large bodies of specialized knowledge and skill involved in their performance. But where does philosophy come in?

Most of us trained in analytic philosophy think of a philosophical topic as one that can be formulated in terms of a small number of familiar questions: what are the nature and limitations of knowledge in this area? What ethical or normative problems does this area raise? What kinds of conceptual issues need to be addressed before we can discuss problems in this area clearly and intelligently? Are there metaphysical issues raised by this area -- special kinds of things that need special philosophical attention? Does "technology" support this kind of analytical approach?

We might choose to pursue a philosophy of technology in an especially minimalist (and somewhat Aristotelian) way, along these lines:

- Human beings have needs and desires that require material objects for their satisfaction.
- Human beings engage in practical activity to satisfy their needs and desires.
- Intelligent beings often seek to utilize and modify their environments so as to satisfy their needs and desires.
- Physical bodies are capable of rudimentary environment modification, which may permit adequate satisfaction of needs and desires in propitious environments (dolphins).
- Intelligent beings often seek to develop "tools" to extend the powers of their bodies to engage in environment modification.
- The use of tools produces benefits and harms for self and others, which raises ethical issues.

Now we can introduce the idea of the accumulation of knowledge ("science"):

- Human beings have the capacity to learn how the world around them works, and they can learn the causal properties of materials and natural entities.
- Knowledge of causal properties permits intelligent intervention in the world.
- Gaining scientific knowledge of the world creates the possibility of the invention of knowledge-based artifacts (instruments, tools, weapons).
- And history suggests we need to add a few Hobbesian premises:
- Human beings often find themselves in conflict with other agents for resources supporting the satisfaction of their needs and desires.
- Intelligent beings seek to develop tools (weapons) to extend the powers of their bodies to engage in successful conflict with other agents.

Finally, history seems to make it clear that tools, machines, and weapons are not purely individual products; rather, social circumstances and social conflict influence the development of the specific kinds of tools, machines, and weapons that are created in a particular historical setting.

The idea of technology can now be fitted into the premises identified here. Technology is the sum of a set of tools, machines, and practical skills available at a given time in a given culture through which needs and interests are satisfied and the dialectic of power and conflict furthered.

This treatment suggests several leading questions for a philosophy of technology:

1. How does technology relate to human nature and human needs?
2. How does technology relate to intelligence and creativity?
3. How does technology relate to scientific knowledge?
4. How does technology fit into the logic of warfare?
5. How does technology fit into the dialectic of social control among groups?
6. How does technology relate to the social, historical, and cultural environment?

7. Is the process of technology change determined by the technical characteristics of the technology?
8. How does technology relate to issues of justice and morality?

Here are a few important contributions to several of these topics.

Lynn White's *Medieval Technology and Social Change* illustrates almost all elements of this configuration. His classic book begins with the dynamics of medieval warfare (the impact of the development of the stirrup on mounted combat); proceeds to food production (the development and social impact of the heavy iron plough); and closes with medieval machines.

Charles Sabel's treatment of industrialization and the creation of powered machinery in *Work and Politics: The Division of Labour in Industry* addresses topic 5; Sabel demonstrates how industrialization and the specific character of mechanization that ensued was a process substantially guided by conflicts of interest between workers and owners, and technologies were selected by owners that reduced the powers of resistance of workers. Sabel and Zeitlin make this argument in greater detail in *World of Possibilities: Flexibility and Mass Production in Western Industrialization*. One of their most basic arguments is the idea that firms are strategic and adaptive as they deal with a current set of business challenges. Rather than an inevitable logic of new technologies and their organizational needs, we see a highly adaptive and selective process in which firms pick and choose among alternatives, often mixing the choices to hedge against failure. They consider carefully a range of possible changes on the horizon, a set of possible strategic adaptations that might be selected; and they frequently hedge their bets by investing in both the old and the new technology. "Economic agents, we found again and again in the course of the seminar's work, do not maximize so much as they strategize" (5).

The logic underlying the idea of technological inevitability (topic 7) goes something like this: a new technology creates a set of reasonably accessible new possibilities for achieving new forms of value: new products, more productive farming techniques, or new ways of satisfying common human needs. Once the technology exists, agents or organizations in society will recognize those new opportunities and will attempt to take advantage of them by investing in the technology and developing it more fully. Some of these attempts will fail, but others will succeed. So over time, the inherent potential of the technology will be realized; the technology will be fully exploited and utilized. And, often enough, the technology will both require and force a new set of social institutions to permit its full utilization; here again, agents will recognize opportunities for gain in the creation of social innovations, and will work towards implementing these social changes.

This view of history doesn't stand up to scrutiny, however. There are many examples of technologies that failed to come to full development (the water mill in the ancient world and the Betamax in the contemporary world). There is nothing inevitable about the way in which a technology will develop -- imposed, perhaps, by the underlying scientific realities of the technology; and there are numerous illustrations of a more complex back-and-forth between social conditions and the development of a technology. So technological determinism is not a credible historical theory.

Thomas Hughes addresses topic 6 in his book, *Human-Built World: How to Think about Technology and Culture*. Here Hughes considers how technology has affected our cultures in the past two centuries. The twentieth-century city, for example, could not have existed without the inventions of electricity, steel buildings, elevators, railroads, and modern waste-treatment technologies. So technology "created" the modern city. But it is also clear that life in the twentieth-century city was transformative for the several generations of rural people who migrated to them. And the literature, art, values, and social consciousness of people in the twentieth century have surely been affected by these new technology systems. Each part of this complex story involves processes that are highly contingent and highly intertwined with social, economic, and political relationships. And the ultimate shape of the technology is the result of decisions and pressures exerted throughout the web of relationships through which the technology took shape. But here is an important point: there is no moment in this story where it is possible to put "technology" on one side and "social context" on the other. Instead, the technology and the society develop together.

Peter Galison's treatment of the simultaneous discovery of the relativity of time measurement by *Einstein and Poincaré in Einstein's Clocks and Poincaré's Maps: Empires of Time* provides a valuable set of insights into topic 3. Galison shows that Einstein's thinking was very much influenced by practical issues in the measurement of time by mechanical devices. This has an interesting corollary: the scientific imagination is sometimes stimulated by technology issues, just as technology solutions are created through imaginative use of new scientific theories.

Topic 8 has produced an entire field of research of its own. The morality of the use of autonomous drones in warfare; the ethical issues raised by CRISPR technology in human embryos; the issues of justice and opportunity created by the digital divide between affluent people and poor people; privacy issues created by ubiquitous facial recognition technology -- all these topics raise important moral and social-justice issues. Here is an interesting thought piece by Michael Lynch in the *Guardian* on the topic of digital privacy (REF). Lynch is the author of *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data*.

So, yes, there is such a thing as the philosophy of technology. But to be a vibrant and intellectually creative field, it needs to be cross-disciplinary, and as interested in the social and historical context of technology as it is the conceptual and normative issues raised by the field.

Section 2

WHAT IS TECHNOLOGY?

We could begin by postulating that technology is ...

the extension of human capacities through the application of scientific knowledge in the design and creation of material artifacts.

Bicycles extended the range and speed of human-powered transportation. Electric lights extended the ability of people to engage in work and leisure after natural light waned. Armored motorized vehicles extended the ability of states to wage war over large territories. Steel plows extended the ability of immigrant farmers to break the sod of the grasslands of the middle west.

The most basic thing we can say is that human beings have material needs, and they are compelled to use tools and artifacts to transform materials provided by nature to satisfy needs. The ensemble of tools, artifacts, practices, and technical knowledge available to a population at a time is its technology. Moreover, human beings are innovative problem solvers. So they are capable of inventing and designing new tools and techniques. This capacity is a primary source of technology change. This is the heart of Marx's insight in *The German Ideology*. [quote]

But technologies are also levers for power and wealth. Control over a technology -- or strong influence over the way in which the technology is developed -- can be a great source of power and wealth for specific groups. And so we need to look closely at the ways in which new technologies are being shaped in ways that serve specific social interests. This is much of what Marx was getting at when he focused on the forces and relations of production as being central to the historical development of a society.

Modern technologies generally require complex human systems in order for them to be broadly implemented. Thomas Hughes documents these complex systems in the case of electric power in *Networks of Power: Electrification in Western Society, 1880-*

1930. Research institutions, engineering firms, municipal governments, and power companies combine to develop and establish the power generation and distribution that basic advances in the understanding of electricity made possible. These institutions aren't guided by a benign optimizing intelligence that produces the optimal implementation for aiding human wellbeing. Rather, they are propelled by private interests, profitability, political competition, and government action. The market plays a role, the demands of the consuming public come in, and the political interests of decision makers and policy mavens are key as well. Technology doesn't direct its own path of development, and neither do the abstract best interests of humanity. (Does the growth of a slime mold colony fit the situation -- locally smart, globally stupid?)

Finally, detailed study of specific technologies -- railroads, steel, chemicals, genetic informatics -- demonstrates a very high degree of contingency in the sequencing of solutions to technical and organizational problems as the technology develops. And these contingencies have significant influence on the outcomes. So technology change is an instance of a path-dependent process.

All of this suggests several important keys for studying the history of various technologies. First, don't make the functionalist assumption that a technology will ultimately develop in a way that is most beneficial to human wellbeing. History is replete with great technical innovations that either quietly disappeared before they could benefit anyone, or were co-opted in ways that primarily benefited elites and power-holders. (The labor-saving water wheel in ancient Rome is a good example.) Second, look for the concrete interests that are at work as the institutional basis and technical solutions for a given technology are chosen. Hughes's discussion of electric power is fundamental. And third, always understand that technology change is a process that demonstrates great contingency and path dependency. So expecting to anticipate the outcome in advance is highly questionable.

Section 3

DEFINING THE PHILOSOPHY OF TECHNOLOGY

The philosophy of technology ought to be an important field within contemporary philosophy, given the centrality of technology in our lives. And yet there is not much of a consensus among philosophers about what the subject of the philosophy of technology actually is. Are we most perplexed by the ethical issues raised by new technological possibilities -- genetic engineering, face recognition, massive databases and straightforward tools for extracting personal information from them? Should we rather ask about the risks created by new technologies -- the risks of technology catastrophe, of unintended health effects, or of further intensification of environmental harms on the planet we inhabit? Should we give special attention to issues of "technology justice" and the inequalities among people that technologies often facilitate, and the forms of power that technology enables for some groups over others? Should we direct our attention to the "existential" issues raised by technology -- the ways that immersion in a technologically intensive world have influenced our development as persons, as intentional and meaning-creating individuals? Are there issues of epistemology, rationality, and creativity that are raised by technology within a social and scientific setting? Should we use this field of philosophy to examine how technology influences human society, and how society influences the development and character of technology? Should we, finally, be concerned that the technology opportunities that confront us encourage an inescapable materialism and a decline of meaningful spiritual or poetic experience?

A useful way of approaching this question is to consider the topics included in the Blackwell handbook, *A Companion to the Philosophy of Technology*, edited by Jan Kyrre Berg Olsen Friis, Stig Andur Pedersen, and Vincent F. Hendricks. The editors and contributors do a good job of attempting to discover philosophical problems in issues raised by technology. The major divisions in this companion include Introduction, History of Technology, Technology and Science, Technology and Philosophy, Technology and Environment, Technology and Politics, Technology and Ethics, and Technology and the Future.

The editors summarize the scope of the field in these terms:

> *The philosophy of technology taken as a whole is an understanding of the consequences of technological impacts relating to the environment, the society and human existence. (Introduction)*

As a definition, however, this attempt falls short. By focusing on "consequences" it leaves unexamined the nature of technology itself, it suggests a unidirectional relationship between technology and human and social life, and it is silent about the normative dimensions of any critical approach to the understanding of technology.

Another useful approach to the topic of how to define the philosophy of technology is Tom Misa's edited collection, *Modernity and Technology*. (Misa's introduction to the volume is available here.) Misa is an historian of technology (he contributes the lead essay on history of technology in the Companion), and he is a particularly astute observer and interpreter of technology in society. His reflections on technology and modernity are especially valuable. Here are a few key ideas:

> *Technologies interact deeply with society and culture, but the interactions involve mutual influence, substantial uncertainty, and historical ambiguity, eliciting resistance, accommodation, acceptance, and even enthusiasm. In an effort to capture these fluid relations, we adopt the notion of co-construction. (3*

This point emphasizes the idea that technology is not a separate historical factor, but rather permeates (and is permeated by) social, cultural, economic, and political realities at every point in time. This is the reality that Misa designates as "co-construction".

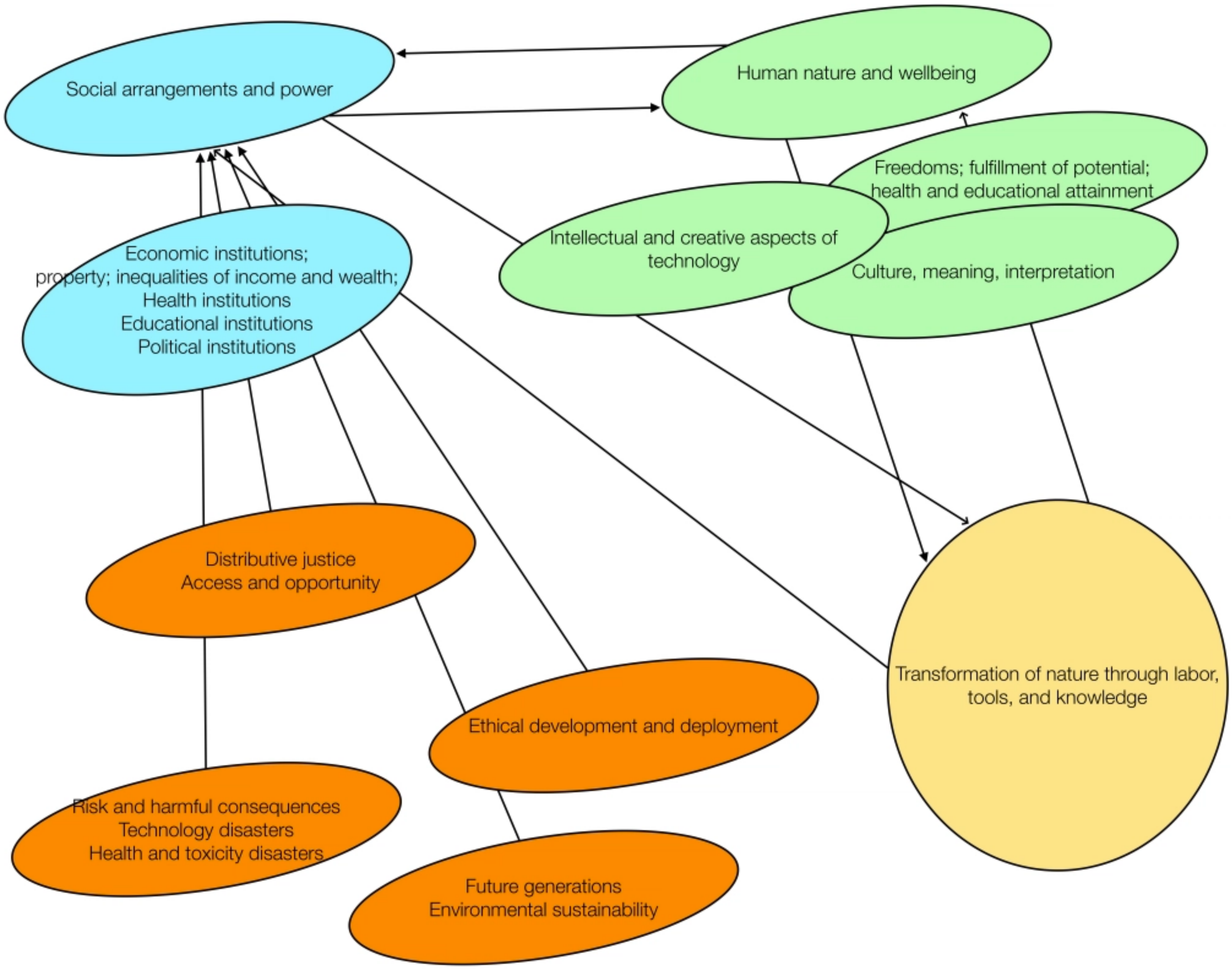
A related insight is Misa's insistence that technology is not one uniform domain that is amenable to analysis and discussion at the purely macro-level. Instead, at any given time the technologies and technological systems available to an epoch are a heterogeneous mix with different characteristics and different ways of influencing human interests. It is necessary, therefore, to address the micro-characteristics of particular technologies rather than "technology in general".

Theorists of modernity frequently conjure a decontextualized image of scientific or technological rationality that has little relation to the complex, messy, collective, problem-solving activities of actual engineers and scientists.... These theorists of modernity invariably posit "technology," where they deal with it at all, as an abstract, unitary, and totalizing entity, and typically counterpose it against traditional formulations (such as lifeworld, self, or focal practices). ... Abstract, reified, and universalistic conceptions of technology obscure the significant differences between birth control and hydrogen bombs, and blind us to the ways different groups and cultures have appropriated the same technology and used it to different ends. To constructively confront technology and modernity, we must look more closely at individual technologies and inquire more carefully into social and cultural processes. (8-9)

And Misa confronts the apparent dichotomy often expressed in technology studies, between technological determinism and social construction of technology:

One can see, of course, that these rival positions are not logically opposed ones. Modern social and cultural formations are technologically shaped; try to think carefully about mobility or interpersonal relations or a rational society without considering the technologies of harbors, railroad stations, roads, telephones, and airports; and the communities of scientists and engineers that make them possible. At the same time, one must understand that technologies, in the modern era as in earlier ones, are socially constructed; they embody varied and even contradictory economic, social, professional, managerial, and military goals. In many ways designers, engineers, managers, financiers, and users of technology all influence the course of technological developments. The development of a technology is contested and controversial as well as constrained and constraining. (10)

It may be that a diagram does a better job of "mapping" the field of the philosophy of technology than a simple definition. Here is a first effort:



The diagram captures the idea that technology is embedded both within the agency, cultures, and values of living human beings during an epoch, and within the social institutions within which human beings function. Human beings and social relations drive the development of technologies, and they are in turn profoundly affected by them. The social institutions include economic institutions (property relations, production and distribution relations), political institutions (institutions of law, policy, and power), and social relations (gender, race, various forms of social inequality). In orange, the diagram represents various kinds of problems of assessment, implementation, development, control, and decision-making that arise in the course of the development and management of technologies, including issues of risk assessment, distribution of burdens and benefits of the effects of technology, and issues concerning future generations and the environment.

A general definition of technology might be framed in these terms: "transformation of nature through labor, tools, and knowledge". And a brief definition of the philosophy of technology, still preliminary, might go along these lines:

The philosophy of technology attempts to uncover the multiple issues raised by "transformation of nature through labor, tools, and knowledge" within the context of large, complex societies. These issues include normative questions, questions of social causation, questions of distributive justice, issues concerning management of risk, and the relationship between technology and human wellbeing.

Section 4

LABOR, FREEDOM, AND TECHNOLOGY

Hegel provided a powerful conception of human beings in the world and a rich conception of freedom. Key to that conception is the idea of self-creation through labor. Hegel had an "aesthetic" conception of labor: human beings confront the raw given of nature and transform it through intelligent effort into things they imagine that will satisfy their needs and desires.

Alexandre Kojève's reading of Hegel is especially clear on Hegel's conception of labor and freedom. This is provided in Kojève's analysis of the Master-Slave section of Hegel's Phenomenology in his *Introduction to the Reading of Hegel*. The key idea is expressed in these terms:

The product of work is the worker's production. It is the realization of his project, of his idea; hence, it is he that is realized in and by this product, and consequently he contemplates himself when he contemplates it.... Therefore, it is by work, and only by work, that man realizes himself objectively as man. (Kojève, Introduction to the Reading of Hegel)

It seems to me that this framework of thought provides an interesting basis for a philosophy of technology as well. We might think of technology as collective and distributed labor, the processes through which human beings collectively transform the world around themselves to better satisfy human needs. Through intelligence and initiative human beings and organizations transform the world around them to create new possibilities for human need satisfaction. Labor and technology are emancipating and self-creating. Labor and technology help to embody the conditions of freedom.

However, this assessment is only one side of the issue. Technologies are created for a range of reasons by a heterogeneous collection of actors: generating profits, buttressing power relations, serving corporate and political interests. It is true that new technologies often serve to extend the powers of the human beings who use them, or to satisfy their needs and wants more fully and efficiently. Profit motives and the market help to ensure that this is true to some extent; technologies and products need to be "desired" if they are to be sold and to generate profits for the businesses that produce them. But given the conflicts of interest that exist in human society, technologies also serve to extend the capacity of some individuals and groups to wield power over others.

This means that there is a dark side to labor and technology as well. There is the labor of un-freedom. Not all labor allows the worker to fulfill him- or herself through free exercise of talents. Instead the wage laborer is regulated by the time clock and the logic of cost reduction. This constitutes Marx's most fundamental critique of capitalism, as a system of alienation and exploitation of the worker as a human being. Here are a few paragraphs on alienated labor from Marx's *Economic and Philosophical Manuscripts*:

The worker becomes all the poorer the more wealth he produces, the more his production increases in power and size. The worker becomes an ever cheaper commodity the more commodities he creates. The devaluation of the world of men is in direct proportion to the increasing value of the world of things. Labor produces not only commodities; it produces itself and the worker as a commodity – and this at the same rate at which it produces commodities in general.

This fact expresses merely that the object which labor produces – labor's product – confronts it as something alien, as a power independent of the producer. The product of labor is labor which has been embodied in an object, which has become material: it is the objectification of labor. Labor's realization is its objectification. Under these economic conditions this realization of labor appears as loss of realization for the workers objectification as loss of the object and bondage to it; appropriation as estrangement, as alienation.

So much does the labor's realization appear as loss of realization that the worker loses realization to the point of starving to death. So much does objectification appear as loss of the object that the worker is robbed of the objects most necessary not only for his life but for his work. Indeed, labor itself becomes an object which he can obtain only with the greatest effort and with the most irregular interruptions. So much does the appropriation of the object ap-

pear as estrangement that the more objects the worker produces the less he can possess and the more he falls under the sway of his product, capital.

All these consequences are implied in the statement that the worker is related to the product of labor as to an alien object. For on this premise it is clear that the more the worker spends himself, the more powerful becomes the alien world of objects which he creates over and against himself, the poorer he himself – his inner world – becomes, the less belongs to him as his own. It is the same in religion. The more man puts into God, the less he retains in himself. The worker puts his life into the object; but now his life no longer belongs to him but to the object. Hence, the greater this activity, the more the worker lacks objects. Whatever the product of his labor is, he is not. Therefore, the greater this product, the less is he himself. The alienation of the worker in his product means not only that his labor becomes an object, an external existence, but that it exists outside him, independently, as something alien to him, and that it becomes a power on its own confronting him. It means that the life which he has conferred on the object confronts him as something hostile and alien.

So does labor fulfill freedom or create alienation? Likewise, does technology emancipate and fulfill us, or does it enthrall and disempower us? Marx's answer to the first question is that it does both, depending on the social relations within which it is defined, managed, and controlled.

It would seem that we can answer the second question for ourselves, in much the same terms. Technology both extends freedom and constricts it. It is indeed true that technology can extend human freedom and realize human capacities. The use of technology and science in agriculture means that only a small percentage of people in advanced countries are farmers, and those who are enjoy a high standard of living compared to peasants of the past. Communication and transportation technologies create new possibilities for education, personal development, and self-expression. The enhancements to economic productivity created by technological advances have permitted a huge increase in the wellbeing of ordinary people in the past century -- a fact that permits us to pursue the things we care about more freely. But new technologies also can be used to control people, to monitor their thoughts and actions, and to wage war against them. More insidiously, new technologies may "alienate" us in new ways -- make us less social, less creative, and less independent of mind and thought.

So it seems clear on its face that technology is both favorable to the expansion of freedom and the exercise of human capacities, and unfavorable. It is the social relations through which technology is exercised and controlled that make the primary difference in which effect is more prominent.

Section 5

KOJÈVE ON FREEDOM

An earlier post highlighted Alexandre Kojève's presentation of Hegel's rich conception of labor, freedom, and human self-creation. This account is contained in Kojève's analysis of the Master-Slave section of Hegel's Phenomenology in Kojève's Introduction to the Reading of Hegel: Lectures on the "Phenomenology of Spirit"; link.

Here are the key passages from Hegel's *Phenomenology* on which Kojève's account depends, from Terry Pinkard's translation in *Georg Wilhelm Friedrich Hegel: The Phenomenology of Spirit*:

Hegel on the Master-Slave relation

195. However, the feeling of absolute power as such, and in the particularities of service, is only dissolution in itself, and, although the fear of the lord is the beginning of wisdom, in that fear consciousness is what it is that is for it itself , but it is not being-for-itself. However, through work, this servile consciousness comes round to itself. In the moment corresponding to desire in the master's consciousness, the aspect of the non-essential relation to the thing seemed to fall to the lot of the servant, as the thing there retained its self-sufficiency. Desire has reserved to itself the pure negating of the object, and, as a result, it has reserved to itself that unmixed feeling for its own self. However, for that reason, this satisfaction is itself only a vanishing, for it lacks the objective aspect, or stable existence. In contrast, work is desire held in check, it is vanishing staved off , or: work cultivates and educates. The negative relation to the object becomes the form of the object; it becomes something that endures because it is just for the laborer himself that the object has self-sufficiency. This negative mediating middle, this formative doing, is at the same time singularity, or the pure being-for-itself of consciousness, which in the work external to it now enters into the element of lasting. Thus, by those means, the working consciousness comes to an intuition of self-sufficient being as its own self.

196. However, what the formative activity means is not only that the serving consciousness as pure being-for-itself becomes, to itself, an exist- ing being within that formative activity. It also has the negative mean- ing of the first moment, that of fear. For in forming the thing, his own negativity, or his being-for-itself, only as a result becomes an object to himself in that he sublates the opposed existing form. However, this objective negative is precisely the alien essence before which he trembled, but now he destroys this alien negative and posits himself as such a negative within the element of continuance. He thereby becomes for himself an existing-being-for-itself . Being-for-itself in the master is to the servant an other, or it is only for him. In fear, being-for-itself is in its own self . In culturally formative activity, being-for-itself becomes for him his own being- for-itself, and he attains the consciousness that he himself is in and for himself. As a result, the form, by being posited as external, becomes to him not something other than himself, for his pure being-for-itself is that very form, which to him therein becomes the truth. Therefore, through this retrieval, he comes to acquire through himself a mind of his own, and he does this precisely in the work in which there had seemed to be only some outsider's mind. – For this reflection, the two moments of fear and service, as well as the moments of culturally formative activity are both necessary, and both are necessary in a universal way. Without the discipline of service and obedience, fear is mired in formality and does not diffuse itself over the conscious actuality of existence. Without culturally formative activity, fear remains inward and mute, and consciousness will not become for it [consciousness] itself. If consciousness engages in formative activity without that first, absolute fear, then it has a mind of its own which is only vanity, for its form, or its negativity, is not negativity in itself , and his formative activity thus cannot to himself give him the consciousness of himself as consciousness of the essence. If he has not been tried and tested by absolute fear but only by a few anxieties, then the negative essence will have remained an externality to himself, and his substance will not have been infected all the way through by it. While not each and every one of the ways in which his natural consciousness was brought to fulfillment was shaken to the core, he is still attached in himself to de-

terminate being. His having a mind of his own is then only stubbornness, a freedom that remains bogged down within the bounds of servility. To the servile consciousness, pure form can as little become the essence as can the pure form – when it is taken as extending itself beyond the singular individual – be a universal culturally formative activity, an absolute concept. Rather, the form is a skill which, while it has dominance over some things, has dominance over neither the universal power nor the entire objective essence. (Hegel, Phenomenology, 115-116)

Kojève's interpretation of Hegel

Here are the primary passages that represent the heart of Kojève's interpretation of this section.

Work, on the other hand, is repressed Desire, an arrested passing phase; or, in other words, it forms-and-educates. Work transforms the World and civilizes, educates, Man, the man who wants to work -- or who must work -- must repress the instinct that drives him "to consume" "immediately" the "raw" object. And the Slave can work for the Master -- that is, for another than himself -- only by repressing his own desires. Hence he transcends himself by working -- or perhaps better, he educates himself, he "cultivates" and "sublimates" his instincts by repressing them. On the other hand, he does not destroy the thing as it is given. He postpones the destruction of the thing by first transforming it through work; he prepares it for consumption -- that is to say, he "forms" it. In his work, he transforms things and transforms himself at the same time: he forms things and the World by transforming himself, by educating himself; and he educates himself, he forms himself, by transforming things and the World, Thus, the negative-or-negating relation to the object becomes a form of this object and gains permanence, precisely because, for the worker, the object has autonomy.... The product of work is the worker's production. It is the realization of his project, of his idea; hence, it is he that is realized in and by this product, and consequently he contemplates himself when he contemplates it.... Therefore, it is by work, and only by work, that man realizes himself objectively as man. Only after producing an artificial object is man himself really and objectively more than and different from a natural being; and only in this real and objective product does he become truly conscious of his subjective human reality. Kojève 24-25

The Master can never detach himself from the World in which he lives, and if this World perishes, he perishes with it. Only the Slave can transcend the given world (which is subjugated by the Master) and not perish. Only the Slave can transform the World that forms him and fixes him in slavery and create a World that he has formed in which he will be free. And the Slave achieves this only through forced and terrified work carried out in the Master's service. To be sure, this work by itself does not free him. But in transforming the World by this work, the Slave transforms himself too, and thus creates the new objective conditions that permit him to take up once more the liberating Fight for recognition that he refused in the beginning for fear of death. And thus in the long run, all slavish work realizes not the Master's will, but the will -- at first unconscious -- of the Slave, who -- finally --succeeds where the Master -- necessarily -- fails. Therefore, it is indeed originally dependent, serving, and slavish Consciousness that in the end realizes and reveals the ideal of autonomous Self-Consciousness and is thus its "truth." Kojève 29-30

However, to understand the edifice of universal history and the process of its construction, one must know the materials that were used to construct it. These materials are men. To know what History is, one must therefore know what Man who realizes it is. Most certainly, man is something quite different from a brick. In the first place, if we want to compare universal history to the construction of an edifice, we must point out that men are not only the bricks that are used in the construction; they are also the masons who build it and the architects who conceive the plan for it, a plan, moreover, which is progressively elaborated during the construction itself. Furthermore, even as "brick," man is essentially different from a material brick: even the human brick changes during the construction, just as the human mason and the human architect do. Nevertheless, there is something in Man, in every man, that makes him suited to participate--passively or actively--in the realization of universal history. At the beginning of this History, which ends finally in absolute Knowledge, there are, so to speak, the necessary and sufficient conditions. And Hegel studies these conditions in the first four chapters of the Phenomenology.

Finally, Man is not only the material, the builder, and the architect of the historical edifice. He is also the one for whom this edifice is constructed: he lives in it, he sees and understands it, he describes and criticizes it. There is a whole category of men who do not actively participate in the historical construction and who are content to live in the constructed edifice and *to talk about it. These men, who live somehow "above the battle," who are content to talk about things that they do not create by their Action, are Intellectuals who produce intellectuals' ideologies, which they take for philosophy (and pass off as such). Hegel describes and criticizes these ideologies in Chapter V. (32-33)*

The central ideas here are --

- Work transforms and educates the worker.
- Work requires the delay of consumption.
- Work transforms the world and the environment.
- The self-creation of the human being through work is essential to his or her reality as a human being.
- By merely directing and commanding work, the master fails to engage in self-creation.
- The master cannot be truly free.
- Human beings create history through their creative labor.
- Human beings create and transform themselves through labor.
- History is human-centered. History is "subject" as well as "object".
- Those who merely think and reflect upon history are sterile and contribute nothing to the course of history.

These comments add up to a substantive theory of the human being in the world -- one that emphasizes creativity, transformation, and self-creation. It stands in stark contrast to the liberal utilitarian view of Adam Smith and Jeremy Bentham of human nature as consumer and rational optimizer of a given set of choices; instead, on Kojève's (and Hegel's) view, the human being becomes fully human through creative engagement with the natural world, through labor.

It is interesting to realize that Kojève was a philosopher, but he was not primarily an academic professor. Instead, he was a high-placed civil servant and statesman in the French state, a man whose thinking and actions were intended to create a new path for France. He is credited with being one of the early theorists of the European Union.

Kojève's account of labor and freedom is, of course, influenced by his own immersion in the writings of the early Marx; so the philosophy of labor, freedom, and self-creation articulated here is neither pure Hegel nor pure Marx. We might say that it is pure Kojève.

Jeff Love's biography of Kojève is also of interest, emphasizing the Russian roots of Kojève's thought; *The Black Circle: A Life of Alexandre Kojève*. Love confirms the importance of the richer theory of human freedom and self-realization that is offered in Kojève's account, and notes a parallel with themes in nineteenth-century Russian literature.

Kojève's critique of self-interest merits renewal in a day when consumer capitalism and the reign of self-interest are hardly in question, either implicitly or explicitly, and where the key precincts of critique have been hobbled by their own reliance on elements of the modern conception of the human being as the free historical individual that have not been sufficiently clarified. Kojève's thought is thus anodyne: far from being "philosophically" mad or the learned jocularity of a jaded, extravagant genius, it expresses a probing inquiry into the nature of human being that returns us to questions that reach down to the roots of the free historical individual. Moreover, it extends a critique of self-interest deeply rooted in Russian thought, and Kojève does so, no doubt with trenchant irony, in the very capital of the modern bourgeoisie decried violently by Dostoevsky in his Winter Notes on Summer Impressions.

(Here is an interesting reflection on Kojève as philosopher by Stanley Rosen; link.)

Section 6

AN EXISTENTIAL PHILOSOPHY OF TECHNOLOGY

March 19, 2020

https://understandingsociety.blogspot.com/2020/03/an-existential-philosophy-of-technology.html

Ours is a technological culture, at least in the quarter of the countries in the world that enjoy a high degree of economic affluence. Cell phones, computers, autonomous vehicles, CT scan machines, communications satellites, nuclear power reactors, artificial DNA, artificial intelligence bots, drone swarms, fiber optic data networks -- we live in an environment that depends unavoidably upon complex, scientifically advanced, and mostly reliable artifacts that go well beyond the comprehension of most consumers and citizens. We often do not understand how they work. But more than that, we do not understand how they affect us in our social, personal, and philosophical lives. We are different kinds of persons than those who came before us, it often seems, because of the sea of technological capabilities in which we swim. We think about our lives differently, and we relate to the social world around us differently.

How can we begin investigating the question of how technology affects the conduct of a "good life"? Is there such a thing as an "existential" philosophy of technology -- that is, having to do with the meaning of the lives of human beings in the concrete historical and technological circumstances in which we now find ourselves? This suggests that we need to consider a particularly deep question: in what ways does advanced technology facilitate the good human life, and in what ways does it frustrate and block the good human life? Does advanced technology facilitate and encourage the development of full human beings, and lives that are lived well, or does it interfere with these outcomes?

We are immediately drawn to a familiar philosophical question, What is a good life, lived well? This has been a central question for philosophers since Aristotle and Epicurus, Kant and Kierkegaard, Sartre and Camus. But let's try to answer it in a paragraph. Let's postulate that there are a handful of characteristics that are associated with a genuinely valuable human life. These might include the individual's realization of a capacity for self-rule, creativity, compassion for others, reflectiveness, and an ability to grow and develop. This suggests that we start from the conception of a full life of freedom and development offered by Amartya Sen in Development as Freedom and the list of capabilities offered by Martha Nussbaum in Creating Capabilities: The Human Development Approach -- capacities for life, health, imagination, emotions, practical reason, affiliation with others, and self-respect. And we might say that a "life lived well" is one in which the person has lived with integrity, justice, and compassion in developing and fulfilling his or her fundamental capacities. Finally, we might say that a society that enables the development of each of these capabilities in all its citizens is a good society.

Now look at the other end of the issue -- what are some of the enhancements to human living that are enabled by modern technologies? There are several obvious candidates. One might say that technology facilitates learning and the acquisition of knowledge; technology can facilitate health (by finding cures and preventions of disease; and by enhancing nutrition, shelter, and other necessities of daily life); technology can facilitate human interaction (through the forms of communication and transportation enabled by modern technology); technology can enhance compassion by acquainting us with the vivid life experiences of others. So technology is sometimes life-enhancing and fulfilling of some of our most fundamental needs and capabilities.

How might Dostoevsky, Dos Passos, Baldwin, or Whitman have adjusted their life plans if confronted by our technological culture? We would hope they would not have been overwhelmed in their imagination and passion for discovering the human in the ordinary by an iPhone, a Twitter feed, and a web browser. We would like to suppose that their insights and talents would have survived and flourished, that poetry, philosophy, and literature would still have emerged, and that compassion and commitment would have found its place even in this alternative world.

But the negative side of technology for human wellbeing is also easy to find. We might say that technology encourages excessive materialism; it draws us away from real interactions with other human beings; it promotes a life consisting of a series of entertaining moments rather than meaningful interactions; and it squelches independence, creativity, and moral focus. So the omnipresence of technologies does not ensure that human beings will live well and fully, by the standards of Aristotle, Epicurus, or Montaigne.

In fact, there is a particularly bleak possibility concerning the lives that advanced everyday technology perhaps encourages: our technological culture encourages us to pursue lives that are primarily oriented towards material satisfaction, entertainment, and toys. This sounds a bit like a form of addiction or substance abuse. We might say that the ambient cultural imperatives of acquiring the latest iPhone, the fastest internet streaming connection, or a Tesla are created by the technological culture that we inhabit, and that these motivations are ultimately unworthy of a fully developed human life. Lucretius, Socrates, and Montaigne would scoff.

It is clear that technology has the power to distort our motives, goals and values. But perhaps with equal justice one might say that this is a life world created by capitalism rather than technology -- a culture that encourages and elicits personal motivations that are "consumerist" and ultimately empty of real human value, a culture that depersonalizes social ties and trivializes human relationships based on trust, loyalty, love, or compassion. This is indeed the critique offered by theorists of the philosophers of the Frankfurt School -- that capitalism depends upon a life world of crass materialism and impoverished social and personal values. And we can say with some exactness how capitalism distorts humanity and culture in its own image: through the machinations of advertising, strategic corporate communications, and the honoring of acquisitiveness and material wealth (link). It is good business to create an environment where people want more and more of the gadgets that technological capitalism can provide.

So what is a solution for people who worry about the shallowness and vapidity of this kind of technological materialism? We might say that an antidote to excessive materialism and technology fetishism is a fairly simple maxim that each person can strive to embrace: aim to identify and pursue the things that genuinely matter in life, not the glittering objects of short-term entertainment and satisfaction. Be temperate, reflective, and purposive in one's life pursuits. Decide what values are of the greatest importance, and make use of technology to further those values, rather than as an end in itself. Let technology be a tool for creativity and commitment, not an end in itself. Be selective and deliberate in one's use of technology, rather than being the hapless consumer of the latest and shiniest. Create a life that matters.

Section 7

BUILDING LARGE TECHNOLOGIES

What is involved in designing, implementing, coordinating, and managing the deployment of a large new technology system in a real social, political, and organizational environment? Here I am thinking of projects like the development of the SAGE early warning system, the Affordable Care Act, or the introduction of nuclear power into the civilian power industry.

Tom Hughes described several such projects in *Rescuing Prometheus: Four Monumental Projects That Changed the Modern World*. Here is how he describes his focus in that book:

Telling the story of this ongoing creation since 1945 carries us into a human-built world far more complex than that populated earlier by heroic inventors such as Thomas Edison and by firms such as the Ford Motor Company. Post-World War II cultural history of technology and science introduces us to system builders and the military-industrial-university complex. Our focus will be on massive research and development projects rather than on the invention and development of individual machines, devices, and processes. In short, we shall be dealing with collective creative endeavors that have produced the communications, information, transportation, and defense systems that structure our world and shape the way we live our lives. (3)

The emphasis here is on size, complexity, and multi-dimensionality. The projects that Hughes describes include the SAGE air defense system, the Atlas ICBM, Boston's Central Artery/Tunnel project, and the development of ARPANET. Here is an encapsulated description of the SAGE process:

The history of the SAGE Project contains a number of features that became commonplace in the development of large-scale technologies. Transdisciplinary committees, summer study groups, mission-oriented laboratories, government agencies, private corporations, and systems-engineering organizations were involved in the creation of SAGE. More than providing an example of system building from heterogeneous technical and organizational compo-

nents, the project showed the world how a digital computer could function as a real-time information-processing center for a complex command and control system. SAGE demonstrated that computers could be more than arithmetic calculators, that they could function as automated control centers for industrial as well as military processes. (16)

Mega-projects like these require coordinated efforts in multiple areas -- technical and engineering challenges, business and financial issues, regulatory issues, and numerous other areas where innovation, discovery, and implementation are required. In order to be successful, the organization needs to make realistic judgments about questions for which there can be no certainty -- the future development of technology, the needs and preferences of future businesses and consumers, and the pricing structure that will exist for the goods and services of the industry in the future. And because circumstances change over time, the process needs to be able to adapt to important new elements in the planning environment.

There are multiple dimensions of projects like these. There is the problem of establishing the fundamental specifications of the project -- capacity, quality, functionality. There is the problem of coordinating the efforts of a very large team of geographically dispersed scientists and engineers, whose work is deployed across various parts of the problem. There is the problem of fitting the cost and scope of the project into the budgetary envelope that exists for it. And there is the problem of adapting to changing circumstances during the period of development and implementation -- new technology choices, new economic circumstances, significant changes in demand or social need for the product, large shifts in the costs of inputs into the technology. Obstacles in any of these diverse areas can lead to impairment or failure of the project.

Most of the cases mentioned here involve engineering projects sponsored by the government or the military. And the complexities of these cases are instructive. But there are equally complex cases that are implemented in a private corporate environment -- for example, the development of next-generation space vehicles by SpaceX. And the same issues of planning, coordination, and oversight arise in the private sector as well.

The most obvious thing to note in projects like these -- and many other contemporary projects of similar scope -- is that they require large teams of people with widely different areas of expertise and an ability to collaborate across disciplines. So a key part of leadership and management is to solve the problem of securing coordination around an overall plan across the numerous groups; updating plans in face of changing circumstances; and ensuring that the work products of the several groups are compatible with each other. Moreover, there is the perennial challenge of creating arrangements and incentives in the work environment -- laboratory, design office, budget division, logistics planning -- that stimulate the participants to high-level creativity and achievement.

This topic is of interest for practical reasons -- as a society we need to be confident in the effectiveness and responsiveness of the planning and development that goes into large projects like these. But it is also of interest for a deeper reason: the challenge of attributing rational planning and action to a very large and distributed organization at all. When an individual scientist or engineer leads a laboratory focused on a particular set of research problems, it is possible for that individual (with assistance from the program and lab managers hired for the effort) to keep the important scientific and logistical details in mind. It is an individual effort. But the projects described here are sufficiently complex that there is no individual leader who has the whole plan in mind. Instead, the "organizational intentionality" is embodied in the working committees, communications processes, and assessment mechanisms that have been established.

It is interesting to consider how students, both undergraduate and graduate, can come to have a better appreciation of the organizational challenges raised by large projects like these. Almost by definition, study of these problem areas in a traditional university curriculum proceeds from the point of view of a specialized discipline -- accounting, electrical engineering, environmental policy. But the view provided from a discipline is insufficient to give the student a rich understanding of the complexity of the real-world problems associated with projects like these. It is tempting to think that advanced courses for engineering and management students could be devised making extensive use of detailed case studies as well as simulation tools that would allow students to gain a more adequate understanding of what is needed to organize and implement a large new system. And interestingly enough, this is a place where the skills of humanists and social scientists are perhaps even more essential than the expertise of technology and management specialists. Historians and sociologists have a great deal to add to a student's understanding of these complex, messy processes.

Section 8

TECHNICAL KNOWLEDGE

There is a kind of knowledge in an advanced mechanical society that doesn't get much attention from philosophers of science and sociologists of science, but it is critical for keeping the whole thing running. I'm thinking here of the knowledge possessed by skilled technicians and fixers -- the people who show up when a complicated piece of equipment starts behaving badly. You can think of elevator technicians, millwrights, aircraft maintenance specialists, network technicians, and locksmiths.

I had an interesting conversation along these lines on the hotel shuttle at the Beijing airport recently. Tim was traveling from Milwaukee to someplace he described as being on the Russian-Mongolian border where there was a mine with a malfunctioning piece of heavy equipment provided by his US company. He expects to be on site for two months, and knows that whatever problems he encounters, they won't be in the users' manual.

This trip is routine for Tim. His company's equipment is used in mines all over the world, from Sweden to India to Brazil. And he is routinely dispatched with his 80-pound duffel, his hard hat, and a few essentials to try to correct the problem.

I said to him, you probably run into problems that don't have a ready solution in the handbook. He said in some amazement, "none of the problems I deal with have textbook solutions. You have to make do with what you find on the ground and nothing is routine." I also asked about the engineering staff back in Wisconsin. "Nice guys, but they've never spent any time in the field and they don't take any feedback from us about how the equipment is failing." He referred to the redesign of a heavy machine part a few years ago. The redesign changed the geometry and the moment arm, and it's caused problems ever since. "I tell them what's happening, and they say it works fine on paper. Ha! The blueprints have to be changed, but nothing ever happens."

I would trust Tim to fix the machinery in my gold mine, if I had one. And it seems that he, and thousands of others like him, have a detailed and practical kind of knowledge about the machine and its functioning in a real environment that doesn't get captured in an engineering curriculum. It is practical knowledge: "If you run into this kind of malfunction, try replacing the thingamajig and rebalance the whatnot." It's also a creative problem-solving kind of knowledge: "Given lots of experience with this kind of machine and these kinds of failures, maybe we could try X." And it appears that it is a cryptic, non-formalized kind of knowledge. The company and the mine owners depend crucially on knowledge in Tim's head and hands that can only be reproduced by another skilled fixer being trained by Tim.

In philosophy we have a few distinctions that seem to capture some aspects of this kind of knowledge: "knowing that" versus "knowing how", *epistime* versus *techne*, formal knowledge versus tacit knowledge. Michael Polanyi incorporated some of these distinctions into his theory of science in *Personal Knowledge: Towards a Post-Critical Philosophy* sixty years ago, but I'm not aware of developments since then.

In sociology and anthropology there has also been some beginning of work on the role that this kind of tacit or non-formalized knowledge plays in the modern technological system. In the early 1980s Xerox Parc commissioned an effort at business anthropology that studied the work practices of Xerox copier repairmen. This was part of a knowledge process called Eureka. The repairmen drive around with vans full of manuals on various models of copier. But it turns out that the bulk of their work depends on shared practical knowledge within the group of repairers at the time. Phone calls are made, interventions are tried, copy machines come back into service. But the manuals are never part of the process.

One of Tim's points seems entirely valid: it is a serious mistake for a company to create a system where engineers design things without ever dealing with their machines in the field. This feedback loop seems critical. The engineers lack access to the tacit technical knowledge that would be gained by practical immersion.

Chuck Sabel's research on machinists in Italy falls in this category of investigation. Interestingly, he too found some of the dissonance Tim reported between the university-educated engineers and the working fixers who actually interfaced with the machines (*Work and Politics: The Division of Labour in Industry*). Another scholar who takes this kind of concrete knowledge seriously is historian of technology Phil Scranton in *Endless Novelty* and also in his research on jet engines.

It seems that there is an opportunity for a very interesting kind of micro research in the sociology of knowledge here: identify a specific technical repair community and interview them in detail to discover what they know, how they know, and how it all works. This knowledge system is difficult to categorize but seems crucial for an advanced technological society.

Section 9

TACIT KNOWLEDGE

Scientist and philosopher Michael Polanyi introduced the idea of "tacit knowledge" in his 1958 book, Personal Knowledge: Towards a Post-Critical Philosophy. The book was presented as a critique of the positivist conception of scientific knowledge and the idea of knowledge as a system of logical statements. Polanyi was trained as a physician before World War I; worked as a research chemist between the wars; and found his voice as a philosopher of science subsequently. (Michael Polanyi is the brother of Karl Polanyi.)

This is an interesting concept, and one that captures an important dimension of knowledge that is absent in most philosophical treatments of epistemology ("knowledge is a system of true justified beliefs"). The simple idea is that there are domains of knowledge that are not represented propositionally or as a system of statements, but are rather somehow embodied in the knower's cognitive system in a non-propositional form. This aspect of knowledge is more analogous to "knowing how" than "knowing that". Polanyi gives the example of a physician in training learning to "read" an x-ray. What is first perceived simply as an unintelligible alternation of light and dark areas, eventually is perceived by the experienced radiologist as a picture of a lung with a tumor. So the physician has somehow acquired a set of perceptual and conceptual skills that cannot be precisely codified but that permit him/her to gain a much more knowledgeable understanding of the patient's hidden disease than the novice. (This seems to be part of what Malcolm Gladwell is getting at in *Blink: The Power of Thinking Without Thinking*.)

It is unremarkable to observe that many aspects of skilled performance depend on "knowledge" that cannot be articulated as a set of statements or rules (link). The basketball guard's ability to weave through defenders and find his way to the basket reflects a complex set of representations of the court, the defenders, and probable behaviors of others that can't be codified. So it seems fairly straightforward to conclude that human cognition incorporates representations and knowledge that do not take the form of explicit systems of statements. Rather, these areas of knowledge are more "intuitive"; they are more akin to "body knowledge." But they are nonetheless cognitive; they depend on experience, they can be criticized and corrected, and they are representational of the world. (Talk to a skilled athlete about a complex task like finding a shot on goal in hockey or beating the defender to the basket, and you will be struck by the degree of intuition, gestalt, and realism that is invoked. And the same is true if you talk to an experienced labor organizer or a police detective.)

What is striking about Polanyi's position in *Personal Knowledge* is that he shows that these forms of practical knowledge do not pertain solely to physical skills like winemaking, playing basketball, or piloting a tug boat. Instead, they extend deeply into the enterprise of scientific knowledge itself. An experimental chemist or physicist has an uncodified ability to interpret instruments, evaluate complex devices, or recognize unexpected results that is the result of experience and training and cannot be reduced to a recipe or algorithm. And the difference between a gifted sociologist and a pedestrian one is his/her ability to probe a set of social facts for the underlying patterns, mechanisms, or questions that they suggest. These are ideas that return again in post-positivist philosophers and historians of science like Thomas Kuhn, Norwood Hanson, and Peter Galison; each of these scholars emphasizes the tight relationship that exists in scientific education and imagination between scientific practices, scientific instruments, and scientific knowledge. Galison's *Einstein's Clocks, Poincare's Maps: Empires of Time* is particularly interesting.

This recognition opens up a number of areas for empirical and theoretical research. The cognitive psychologist can ask a series of questions about how this kind of knowledge is represented in the knower's cognitive system. Does the basketball guard have a "cognitive map" of the court that he updates regularly? (This is a subject for non-human cognitive psychology as well as human cognition.) And the phenomenological sociologist like Erving Goffman can probe in detail the forms of tacit knowledge that the skilled intellectual performer is making use of -- the lawyer or accountant, or the research scientist (link).

This line of thought converges to some extent with arguments that Hubert Dreyfus advanced in What computers can't do: A critique of artificial reason in 1972. (Here is his update of his position; *What Computers Still Can't Do: A Critique of Artificial Reason*.) Dreyfus was fundamentally critical of artificial intelligence research in the 1960s, and the strategy of attempting to codify expert knowledge in the form of a set of rules that could then be implemented as computational algorithms. His position was a phenomenological one; basically, he took issue with the idea that cognitive competences like chess-playing, problem-solving, or pattern recognition could be reduced to a set of precise and separate rules and statements. Instead, there is a holistic aspect in ordinary practical knowledge that cannot be reduced to a set of discrete algorithms.

Of special interest for *Understanding Society* is the question of whether ordinary people have "tacit social knowledge" of the social world they inhabit (link, link, link). What is involved in the competence a person demonstrates when he/she successfully navigates a formal dinner or a contentious union-hall argument? Can the knowledge that the competent social participant has about expected behavior from particular individuals be represented as a sum of propositional beliefs? Or, more plausibly, is this a good example of tacit knowledge, more akin to a rough map of a terrain than a codified set of statements? What about failures of tacit knowledge -- clumsiness (link)? And what is the system of social cognition through which ordinary social knowers gain these cognitive competences and update them through subsequent experience? All of this makes me think that we need richer models of mental life and competence than we currently possess (link).

Chapter 3

TECHNOLOGY AND HISTORY

Section 1

TECHNOLOGY AND SOCIETY

Technology is sometimes thought of as a domain with a logic of its own -- an inevitable trend towards the development of the most efficient artifacts, given the potential represented by a novel scientific or technical insight. The most important shift that has occurred in the ways in which historians conceptualize the history of technology in the past thirty years is the clear recognition that technology is a social product, all the way down. And, as a corollary, historians of technology have increasingly come to recognize the deep contingency that characterizes the development of specific instances or families of technologies.

Thomas Hughes is one of the most important and prolific historians of technology of his generation. His most recent book, *Human-Built World: How to Think about Technology and Culture*, is well worth reading. It looks at "technology" from a very broad perspective and asks how this dimension of civilization has affected our cultures in the past two centuries. The twentieth-century city, for example, could not have existed without the inventions of electricity, steel buildings, elevators, railroads, and modern waste-treatment technologies. So technology "created" the modern city. But it is also clear that life in the twentieth-century city was transformative for the several generations of rural people who migrated to them. And the literature, art, values, and social consciousness of people in the twentieth century have surely been affected by these new technology systems.

This level of analysis stands at the most generic perspective: how does technology influence culture? (And perhaps, how does culture influence technology?) What Hughes has demonstrated in so much of his work, though, is the fact that the most interesting questions about the "technology-society" interface can be framed at a much more disaggregated level. Consider some of the connections he suggests in his earlier book on the history of electric power (*Networks of Power: Electrification in Western Society, 1880-1930*):

- Invention (by individuals with a very specific educational and cultural background)
- Concrete development of the artifacts within a laboratory (involving specific social relationships among various experts and workers)
- "Selling" the innovation to municipal authorities (for lighting and traction) and to industrial capitalists (for power)
- Finding investors and sources of finance for large capital investments in electricity
- Building out the infrastructure for delivery of electric power
- Government regulation of industry practices
- Development of an extended research capability addressing technology problems

Each part of this complex story involves processes that are highly contingent and highly intertwined with social, economic, and political relationships. And the ultimate shape of the technology is the result of decisions and pressures exerted throughout the web of relationships through which the technology took shape. But here is an important point: there is no moment in this story where it is possible to put "technology" on one side and "social context" on the other. Instead, the technology and the society develop together.

Hughes also explores some of the ways in which the culture of the machine has influenced architecture, art, and literature. He discusses photography by Charles Sheeler (whose famous series on the Rouge plant defined an industrial aesthetic), artists Carl Grossberg and Marcel Duchamp, and architects such as Peter Behren. The central theme here is the idea that industrial-technological developments caused significant cultural change in Europe and America. Hughes's examples are mostly drawn from "high" culture; but historians of popular culture too have focused on the impact of technologies such as the railroad, the automobile, or the cigarette on American popular culture. See Deborah Clarke's *Driving Women: Fiction and Automobile Culture in Twentieth-Century America* for a discussion of the effect of automotive culture. And Pam Pennock's examination of the effects of alcohol and tobacco advertising on American culture in *Advertising Sin And Sickness: The Politics of Alcohol And Tobacco Marketing, 1950-1990* is also relevant.

Hughes doesn't consider here the other line of influence that is possible between culture and technology: how prevailing aesthetic and cultural preferences influence the development of a technology. This has been an important theme in the line of interpretation referred to as the "social construction of technology" (SCOT). Wiebe Bijker makes the case for the social construction of mundane technologies such as bicycles in *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change*. And automobile historian Gijs Moms argues in *The Electric Vehicle: Technology and Expectations in the Automobile Age* that the choice between electric and internal combustion vehicles in the early twentieth century turned on aesthetic and lifestyle preferences rather than technical or economic efficiency. This too is a more disaggregated approach to the question. It proceeds on the idea that we can learn a great deal by examining the "micro" processes in culture and society that influence the development of a technology.

It seems to me that the conceptual framework of "assemblages theory" would be useful in discussing the history of technology. (See Manuel DeLanda's *A New Philosophy of Society: Assemblage Theory And Social Complexity* for a review of the theory.) The framework is useful here because technology is a social phenomenon that extends from one's own kitchen and household to the cities of Chicago or Berlin, to the global internet and the international system of manufacturing and design. And similar processes of shaping and conditioning occur at the micro, meso, and macro levels. In other words -- perhaps we can understand "technology" at the molar level, as a complex composition of activities and processes at many levels closer to the socially constructed individual. And the value-added provided by the sociology and history of technology is precisely this: to shed light on the mechanisms at work at all levels that have an influence on the aggregate direction and shape of the resulting technology.

Since we're thinking about "technology and culture" -- it's worth noting that *Technology and Culture* is the world's leading journal for the history of technology, emanating from the Society for the History of Technology (SHOT, established in 1958).

The journal has played a significant role in the definition of the discipline over the past thirty years or so and is an outstanding source for anyone interested in the questions posed here.

Section 2

TECHNOLOGICAL INEVITABILITY

Some historians imagine that new technologies force other kinds of social changes, or even that a given technology creates a more or less inevitable process of development in society. Marx is sometimes thought to offer such a theory: "the forces of production create changes in the relations of production." This view is referred to as technological determinism.

The logic underlying this interpretation of history goes something like this: a new technology creates a set of reasonably accessible new possibilities for achieving new forms of value: new products, more productive farming techniques, or new ways of satisfying common human needs. Once the technology exists, agents or organizations in society will recognize those new opportunities and will attempt to take advantage of them by investing in the technology and developing it more fully. Some of these attempts will fail, but others will succeed. So over time, the inherent potential of the technology will be realized; the technology will be fully exploited and utilized. And, often enough, the technology will both require and force a new set of social institutions to permit its full utilization; here again, agents will recognize opportunities for gain in the creation of social innovations, and will work towards implementing these social changes.

This view of history doesn't stand up to scrutiny, however. There are many examples of technologies that failed to come to full development (the water mill in the ancient world and the Betamax in the contemporary world). There is nothing inevitable about the way in which a technology will develop -- imposed, perhaps, by the underlying scientific realities of the technology; and there are numerous illustrations of a more complex back-and-forth between social conditions and the development of a technology. So technological determinism is not a credible historical theory.

For more credible interpretations of the relationships that exist between technology and historical change, we can consider the work of some very gifted historians.

An example of such a study of social-technological development is offered in William Cronon's fascinating history of Chicago, *Nature's Metropolis: Chicago and the Great West*. Cronon tracks the effects that the extension of the rail network had on the city of Chicago and the region surrounding it into Illinois, Wisconsin, Iowa, and Michigan. Cheap, reliable rail transport between Chicago and New York created large markets for grain and beef. This gave incentives to farmers and traders to organize their activities in such a way as to take advantage of the profits newly available in these markets. But Cronon points out that transportation by railroad of large volumes of grain required a reorganization of the market institutions that were used: the establishment of grain elevators along the rail lines, the establishment of a grading system for qualities of grain being sold by farmers to elevator operators, and the establishment of a futures market for grain and beef. And he observes that entrepreneurs recognized the gains that could be achieved by developing these institutions and carried them forward. So the technology "needed" a reorganization of the market for grain; entrepreneurs recognized an opportunity for profits in achieving this reorganization; and the necessary social innovations occurred.

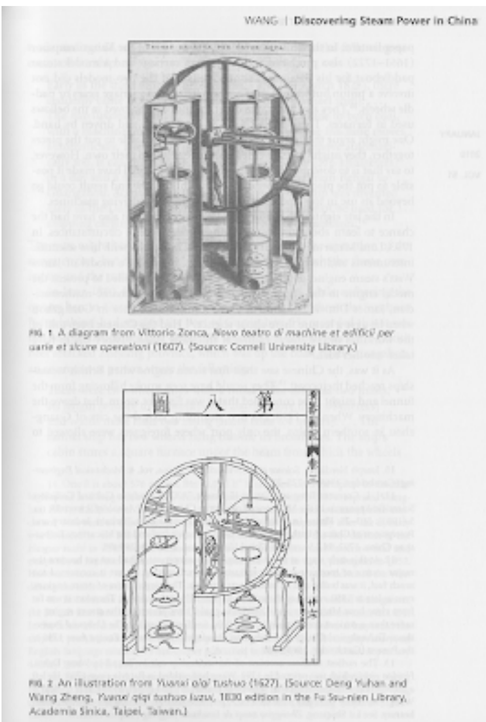
This makes the business and ecological transformation of Chicago's region sound inevitable, given the extension of the rail system into Chicago. However, contingency comes into this story from numerous angles. The build-out of the American rail network was itself a highly contingent matter; the major east-west lines could have been placed in numerous alternative routes, including a network that would have made St. Louis the major rail nexus in the center of the country. Second, the policy environment within which the American rail network developed represented another major form of contingency. As Frank Dobbin demonstrates, England, France, and the United States possessed very different "policy cultures", and these differences created substantial differences in the way in which the basic technology of the railroad was exploited in the three national settings (*Forging Industrial Policy: The United States, Britain, and France in the Railway Age*). And third, there are multiple social solutions that would work roughly as well as the institutions of the grain elevator and the futures market for solving the business challenges of mass transport and marketing of grain. The solutions that emerged in Chicago were therefore contingent as well.

Cronon and Dobbin illustrate several different aspects of technological causes and technological contingency in their accounts of the railroad. But for an author who takes the contingency of technology change even more seriously, see Thomas Hughes' history of electric power in *Networks of Power: Electrification in Western Society, 1880-1930*. Hughes demonstrates that the basic scientific discoveries associated with the technology of electric power had multiple possible realizations in machines and devices, and that the social constraints that existed in different national scientific-technical-political cultures led to substantial and fundamental differences in the development of the technology. The celebrated contest between alternating and direct current as the foundation of power transmission is one such example, but there are hundreds of other examples of choice that can be uncovered in the history of this fundamental modern technology. And, most pointedly, Hughes illustrates the way in which different municipal political requirements in England, France, and the United States led to substantially different implementations of the power stations and transmission networks in the three national contexts.

These examples illustrate several fundamental points about the role of technology in historical change. A new technology creates new opportunities for power, wealth, efficiency, or productivity; so a new technology can be a powerful force for social and economic change. Governments, farmers, entrepreneurs, and corporations have a complicated set of incentives that lead them to consider developing the new technology. So new technologies certainly function as effective historical causes. The development of a technology, however, introduces deeply significant elements of contingency. The term "path-dependency" is an accurate description of the process of the development of a major technology. Third, a technology is both influenced by social factors in the society in which it is developed, and also influences the future direction of social factors in the society. Thus technology is both cause and effect of social change. And, finally, it is evident that the study of the history of technology is inevitably a study of social processes and institutions as much as it is a study of machines and inventions. Technology is a social product, shaped by the needs and powers that exist in society as much as it is shaped by scientific imagination and discovery.

Section 3

TRANSMITTING TECHNOLOGY



November 12, 2010

https://understandingsociety.blogspot.com/2010/11/transmitting-technology.html

How do large technological advances cross cultural and civilizational boundaries? The puzzle is this: large technologies are not simply cool new devices, but rather complex systems of scientific knowledge, engineering traditions, production processes, and modes of technical communication. So transfer of technology is not simply a matter of conveying the approximate specifications of the device; it requires the creation of a research and development infrastructure that is largely analogous to the original process of invention and development. Inventors, scientists, universities, research centers, and skilled workers need to build a local understanding of the way the technology works and how to solve the difficult problems of material and technical implementation.

Take inertial guidance systems for missiles, described in fascinating detail by Donald MacKenzie in *Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance*. The process MacKenzie describes of discovery and development of inertial guidance was a highly complex and secretive one, with multiple areas of scientific and engineering research solving a series of difficult technical problems.

Now do a bit of counterfactual history and imagine that some country -- say, Burma -- had developed powerful rocket engines in the 1950s but did not have a workable guidance technology; and suppose the US and USSR had succeeded in keeping the development of inertial navigation systems and the underlying science secret. Finally, suppose that Burmese agents had managed to gain a superficial description of inertial navigation: "It is a self-contained device that tracks acceleration and therefore permits constant updating of current location; and it uses ultra-high precision gyroscopes." Would this be enough of a leak to permit rapid adoption of inertial navigation in the Burmese missile program? Probably not; the technical obstacles faced in the original development process would have to be solved again, and this means a long process of knowledge building and institution building. For example, MacKenzie describes the knotty problem posed to this technology by the fact of slight variations in the earth's gravitational field over the surface of the globe; if uncorrected, these variations would be coded as acceleration by the instrument and would lead to significant navigational errors. The solution to this problem involved creating a detailed mapping of the earth's gravitational field.

This is a hypothetical case. But Hsien-Chun Wang describes an equally fascinating but real case in a recent article in *Technology & Culture*, "Discovering Steam Power in China, 1840s-1860s" (link). There was essentially no knowledge of steam power in Chinese science in the mid-Qing (early nineteenth century). The First Opium War (1839-1842) provided a rude announcement of the technology, in the form of powerful steam-driven warships on the coast and rivers of eastern China. Chinese officials and military officers recognized the threat represented by Western military-industrial technology, but it was another 25 years before Chinese industry was in a position to build a steam-powered ship. So what were the obstacles standing in front of China's steam revolution?

Wang focuses on two key obstacles in mid-Qing industry and technology: the role of technical drawings as a medium for transmitting specifications for complex machines from designer to skilled workers; and the absence in nineteenth-century

China of a machine tool technology. Technical drawings were an essential medium of communication in the European industrial system, conveying precise specifications of parts and machines to the workers and tools who would fabricate them. And machine tools (lathes, planes, stamping machines, cutting machines, etc.) provided the tools necessary to fabricate high-precision metal parts and components. (Merritt Roe Smith describes aspects of both these stories in his account of the U.S. arms industry in the early nineteenth century; *Harpers Ferry Armory and New Technology*.) According to Wang, the Chinese technical culture had developed models rather than drawings to convey how a machine works; and the intricate small machines that certainly were a part of Chinese technical culture depended on artisanal skill rather than precision tooling of interchangeable parts.

So communicating the technical details of a complex machine and creating the fabrication infrastructure needed to produce the machine were two important obstacles for rapid transfer of steam technology from Western Europe to Qing China. But perhaps a more fundamental obstacle emerges as well: the fact that Chinese technical and scientific culture was as yet simply unready to "see" the way that steam power worked in the first place. When steam warships arrived, acute Chinese observers saw smoke and fire, and they saw motion. But they did not see "steam-driven traction", or the translation of kinetic energy into rotational work. (This is evident also in the drawing of the treadmill water pump above; the maker of the drawing clearly did not perceive from the Italian drawing how the motion of the treadmill was translated into the vertical pumping action.) Wang quotes a description from an observer in Guangzhou in 1828:

> *Early in the third month ... there suddenly came from Bengal a huo lunchuan [fire-wheel ship] .... The huo lunchuan has an empty copper cylinder inside to burn coal, with a machine on the top. When the flame is up, the machine moves automatically. The wheels on both sides of the ship move automatically too. (37)*

And another observer wrote in Zhejiang in 1840:

The ship's cabin stores a square furnace under the beam from which the wheels are hung. When the fire is burning in the furnace, the two wheels turn like a fast mill and the ship cruises as fast as if it is flying, regardless of the wind's direction. (37-38)

The give-away here is the word "automatically"; plainly these observers had not assimilated a causal process linking the production of heat (fire) to mechanical motion (the rotation of the paddle wheels). Instead, the two circumstances are described as parallel rather than causal.

So the fundamental motive force of steam was not cognitively accessible at this point, even through direct observation. By contrast, the marine utility of paddle-wheel-driven warships was quickly assimilated. Chinese commanders adapted what they observed in the European naval forces (powerful paddlewheels that made sails unnecessary) to an existing technology (human- or ox-driven paddle-wheels), and large "wheel-boats" saw action as early as 1842 on Suzhou Creek (40).

Wang notes that several Chinese inventors did in fact succeed in discerning the mechanism associated with steam power by the 1840s. Ding Gongchen succeeded in fabricating a model steam rail engine 61 centimeters long and a 134-centimeter model paddlewheel steamboat; so he clearly understood the basic mechanism at this point. And Zheng Fuguang appears to have mastered the basic concept as well. But here is Wang's summary:

Ding's efforts show that despite the circulating writings of a few experimenters, the steam engine remained a novelty, which was difficult to understand and probably impossible to reproduce. Interested parties were discussing it, however, but attempted to grasp it in terms of their indigenous expertise alone rather than more broadly understanding the new Western technology. (45)

In 1861, during the Taiping Rebellion, a senior military commander Zeng Guofan created an arsenal in Anqing for ammunition, and also set about to create the capacity to build steam-powered ships. With the assistance of experts Xu Shou and Hua Hengfang, the arsenal produced a partially successful full-scale steamship by 1863, and in 1864 Hua and Xu succeeded in completing a 25-ton steamship, the Huanghu, that was capable of generating 11.5 kilometers per hour. The Chinese-build steamship had arrived.

Here is how Wang summarizes this history of technology adaptation over a 25-year period of time:

The path from the treadmill paddlewheel boat to the Jiangnan arsenal's steamers was a long journey of discovery. Qing officials experimented with the knowledge and skills available to them, and it took time--and trial and error--for them to realize that steamboats were driven by steam, that machine tools were necessary to turn the principle of steam into a workable engine, and that drawings had to be made and read for the technology to be diffused. (53)

So perhaps the short answer to the question posed above about cross-civilizational technology transfer is this: "transfer" looks a lot more like "reinvention" than it does "imitation." It was necessary for Chinese experimenters, officials, and military officers to create a new set of institutions and technical capacities before this apparently simple new technological idea could find its way into Chinese implementations on a large scale.

(The image at the top is one of the most interesting parts of Wang's very interesting paper; it establishes vividly the difficulty of transmitting technologies across different technical cultures. The Italian drawing dates from 1607, and the Chinese copy dates from 1627. As Wang points out, the Chinese version of the drawing is visually highly similar to the Italian original; it is a good copy. And yet it fails to designate any of the technical features of how this treadmill-operated water pump works. The pair of drawings are fascinating to examine in detail.)

Section 4

UNINTENDED CONSEQUENCES

History offers plentiful examples of the phenomenon of unintended consequences -- for example, wars that break out unexpectedly because of actions taken by states to achieve their security, or financial crises that erupt because of steps taken to avert them. (The recent military escalations in Pakistan and India raise the specter of unintended consequences in the form of military conflict between the two states.) But technology development, city planning, and economic development policy all offer examples of the occurrence of unintended consequences deriving from complex plans as well.

Putting the concept schematically -- an actor foresees an objective to be gained or an outcome to be avoided. The actor creates a plan of action designed to achieve the objective or avert the undesired outcome. The plan is based on a theory of the causal and social processes that govern the domain in question and the actions that other parties may take. The plan of action, however, also creates an unforeseen or unintended series of developments that lead to a result that is contrary to the actor's original intentions.

It's worth thinking about this concept a bit. An unintended consequence is different than simply an undesired ***outcome***; a train wreck or a volcano is not an unintended consequence, but rather simply an unfortunate event. Rather, the concept fits into the framework of intention and purposive action. An unintended consequence is a result that came about because of deliberate actions and policies that were set in train at an earlier time -- so an unintended consequence is the result of deliberate action. But the outcome is not one of the goals to which the plan or action was directed; it is "unintended". In other words, analysis of the concept of unintended consequences fits into what we might call the "philosophy of complex action and planning." (Unlikely as this sub-specialty of philosophy might sound, here's a good example of a work in this field by Michael Bratman, *Intention, Plans, and Practical Reason*. Robert Merton wrote about the phenomenon of unintended consequences quite a bit, based on his analysis of the relationships between policy and social science knowledge, in *Social Theory and Social Structure*.)

But there is also an element of paradox in our normal uses of the concept of an unintended consequence -- the suggestion that plans of action often contain elements that work out to defeat them. The very effort to bring about X creates a dynamic that frustrates the achievement of X. This is suggested by the phrase, the "law of unintended consequences." (I think this is what Hegel refers to as the cunning of reason.)

There is an important parallel between unintended and unforeseen consequences, but they are not the same. A harmful outcome may have occurred precisely because because it was unforeseen -- it might have been easily averted if the planner had been aware of it as a possible consequence. An example might be the results of the inadvertent distribution of a contaminant in the packaging of a food product. But it is also possible that an undesired outcome is both unintended but also fully foreseen. An example of this possibility is the decision of state legislators to raise the speed limit to 70 mph. Good and reliable safety statistics make it readily apparent that the accident rate will rise. Nonetheless the officials may reason that the increase in efficiency and convenience more than offsets the harm of the increase in the accident rate. In this case the harmful result is unintended but foreseen. (This is the kind of situation where cost-benefit analysis is brought to bear.)

Is it essential to the idea of unintended consequences that the outcome in question be harmful or undesirable? Or is the category of "beneficial unintended consequence" a coherent one? There does seem to be an implication that the unintended consequence is one that the actor would have avoided if possible, so a beneficial unintended consequence violates this implicature. But I suppose we could imagine a situation like this: a city planner sets out to design a park that will give teenagers a place to play safely, increase the "green" footprint of the city, and draw more families to the central city. Suppose the plan is implemented and each goal is achieved. But it is also observed that the rate of rat infestation in surrounding neighborhoods falls dramatically -- because the park creates habitat for voracious rat predators.

This is an unintended but beneficial consequence. And full knowledge of this dynamic would not lead the planner to revise the plan to remove this feature.

The category of "unintended but foreseen consequences" is easy to handle from the point of view of rational planning. The planner should design the plan so as to minimize avoidable bad consequences; then do a cost-benefit analysis to assess whether the value of the intended consequences outweighs the harms associated with the unintended consequences.

The category of consequences of a plan that are currently unforeseen is more difficult to handle from the point of view of rational decision-making. Good planning requires that the planner make energetic efforts to canvass the consequences the plan may give rise to. But of course it isn't possible to discover all possible consequences of a line of action; so the possibility always exists that there will be persistent unforeseen negative consequences of the plan. The most we can ask, it would seem, is that the planner should exercise due diligence in exploring the most likely collateral consequences of the plan. And we might also want the planner to incorporate some sort of plan for "soft landings" in cases where unforeseen negative consequences do arise.

Finally, is there a "law of unintended consequences", along the lines of something like this:

"No matter how careful one is in estimating the probable consequences of a line of action, there is a high likelihood that the action will produce harmful unanticipated consequences that negate the purpose of the action."

No; this statement might be called "reverse teleology" or negative functionalism, and certainly goes further than empirical experience or logic would support. The problem with this statement is the inclusion of the modifier "high likelihood". Rather, what we can say is this:

"No matter how careful one is in estimating the probable consequences of a line of action, there is the residual possibility that the action will produce harmful unanticipated consequences that negate the purpose of the action."

And this statement amounts to a simple, prudent observation of theoretical modesty: we can't know all the possible results of an action undertaken. Does the possibility that any plan may have unintended harmful consequences imply that we should not act? Certainly not; rather, it implies that we should be as ingenious as possible in trying to anticipate at least the most likely consequences of the contemplated actions. And it suggests the wisdom of action plans that make allowances for soft landings rather than catastrophic failures.

Section 5

PATH DEPENDENCY

Historical outcomes are rarely determined by some set of antecedent conditions. Instead, the outcome that occurs is the result of an extended series of events and conditions, each stage of which is subject to substantial contingency and conjunctures. This is most visible in the occurrence of large, dramatic events like stock market crashes, peasant uprisings, and wars. But it is also true in more mundane occurrences -- scandals, resignations, and insolvencies, for example.

The idea of path dependency is chiefly relevant in application to particular events and happenings. But there are large social structures that display path dependency as well. These examples amount to structuring conditions that persist for a long time and further condition other historical developments. Their emergence is contingent, and other directions were feasible in the early stage. But once the structure is in place it creates its own conditions for persistence. Because the Interstate highway program privileged automobile transport over rail in the 1950s, it is now difficult to create a financially viable passenger rail system in the United States.

Examples of historical developments that display significant path dependency include the choice of major technologies (internal combustion engines versus electric vehicles, alternating current rather than direct current as the long-distance power transmission standard, the QWERTY keyboard versus other more efficient arrangements), major infrastructures (rail rather than canal in the late nineteenth century, personal automobile rather than passenger rail in the mid-twentieth century), and the division of tasks and specialization across adjacent professions (social workers, psychiatrists). In each of these cases there is a contingent first step that could have led in another direction, and then an accumulated resistance to reverting to the other choice (substantial social investment already committed to the first option, substantial development of the knowledge systems required for the first option but not the second, substantial social power invested in individuals and organizations that have a vested interest in maintaining the first option, etc.). (Tom

Hughes refers to these factors as "technological momentum" in *Networks of Power: Electrification in Western Society, 1880-1930*.) It isn't impossible to retrace steps and adopt the alternative solution -- but it is very difficult and uncommon.

James Mahoney has done useful work in explicating this concept in "Path dependence in historical sociology" (link). Here is how he describes his own position:

In this article, I argue that path dependence characterizes specifically those historical sequences in which contingent events set into motion institutional patterns or event chains that have deterministic properties. The identification of path dependence therefore involves both tracing a given outcome back to a particular set of historical events, and showing how these events are themselves contingent occurrences that cannot be explained on the basis of prior historical conditions. (507-508)

Mahoney identifies two basic mechanisms that are associated with path dependency:

First, some path-dependent investigators analyze self-reinforcing sequences characterized by the formation and long-term reproduction of a given institutional pattern. (508)

A second basic type of path-dependent analysis involves the study of reactive sequences. Reactive sequences are chains of temporally ordered and causally connected event. (509)

Here are the three features Mahoney finds to be critical.

I suggest that all path-dependent analyses minimally have three defining features. First, path-dependent analysis involves the study of causal processes that are highly sensitive to events that take place in the early stages of an overall historical sequence. (510)

Second, in a path-dependent sequence, early historical events are contingent occurrences that cannot be explained on the basis of prior events or "initial conditions." (511)

Third, once contingent historical events take place, path-dependent sequences are marked by relatively deterministic causal patterns or what can be thought of as "inertia" - i.e., once *processes are set into motion and begin tracking a particular outcome, these processes tend to stay in motion and continue to track this outcome. (511)*

Mahoney's work here represents a great example of careful, detailed analysis that crosses the categories of methodology, theory, and philosophy. This is how we make progress in the social sciences -- by taking theories and concepts seriously and providing careful, logical analysis of the ways these concepts express important features of the historical process.

The idea of path dependency invokes the intersection of two important intuitions which are in tension with each other. The first is the sense of contingency that study of important historical transitions instills. The second is the idea that explanation involves generalizing across cases. If a case is wholly particular, then it is hard to find anything explanatory in studying it. Careful analysis of the concept of path dependency has the potential of offering a higher-level resolution of these intuitions about the historical process.

Section 6

MARX'S MATERIALISM

Karl Marx was a materialist thinker. But what does this amount to? What is materialism as a way of thinking about historical and social reality? Is materialism an empirical theory, a philosophical theory, or perhaps part of a social-science paradigm?

Here is a statement of Marx's materialism from the *German Ideology*, written in 1845-46:

The premises from which we begin are not arbitrary ones, not dogmas, but real premises from which abstraction can only be made in the imagination. They are the real individuals, their activity and the material conditions under which they live, both those which they find already existing and those produced by their activity. These premises can thus be verified in a purely empirical way.

The first premise of all human history is, of course, the existence of living human individuals. Thus the first fact to be established is the physical organisation of these individuals and their consequent relation to the rest of nature. Of course, we cannot here go either into the actual physical nature of man, or into the natural conditions in which man finds himself – geological, hydrographical, climatic and so on. The writing of history must always set out from these natural bases and their modification in the course of history through the action of men.

Men can be distinguished from animals by consciousness, by religion or anything else you like. They themselves begin to distinguish themselves from animals as soon as they begin to produce their means of subsistence, a step which is conditioned by their physical organisation. By producing their means of subsistence men are indirectly producing their actual material life.

The way in which men produce their means of subsistence depends first of all on the nature of the actual means of subsistence they find in existence and have to reproduce. This mode of *production must not be considered simply as being the production of the physical existence of the individuals. Rather it is a definite form of activity of these individuals, a definite form of expressing their life, a definite mode of life on their part. As individuals express their life, so they are. What they are, therefore, coincides with their production, both with what they produce and with how they produce. The nature of individuals thus depends on the material conditions determining their production.*

This production only makes its appearance with the increase of population. In its turn this presupposes the intercourse [Verkehr] of individuals with one another. The form of this intercourse is again determined by production.

To start, Marx's words here are directed against "philosophy", and the Young Hegelians in particular. So Marx is advocating for a different form of reasoning -- not speculative philosophical reflection, but concrete analysis grounded in knowledge of the circumstances of human life. Marx is saying that we can understand certain important things -- for example, the development of ideas or religion -- by examining the "material" circumstances of life in which they emerge. And what are those circumstances? They are circumstances of material need and human labor: the fact that human beings satisfy their material needs on the basis of the transformation of nature through labor. So what is "material" in this setting is two characteristics: the material needs that human beings have (food, shelter, warmth) and the material-physical properties of the world in which human beings find themselves. Human beings as "producers" -- intelligent transformers of nature through individual and social labor -- this is the fundamental material fact in this passage.

History comes into this account through Marx's reference to the "nature of the actual means of subsistence they find in existence", because this is the social history of technology. Human beings create new tools and techniques over time. So materialism, when applied to human beings, has an inherently historical character; men and women transform the tools and knowledge they use in order to transform nature and satisfy needs. And, as the following paragraphs make clear, the social relations through which production takes place are themselves historical products, in a regular process of change and development. Social relationships are "material" insofar as they are the forms of cooperation through which labor and production take place; central among these material social relationships are the property relations of a given level of society.

These comments focus primarily on the conditions of production as a foundation for materialism. A related line of thought in Marx's writings is the idea of the social relations of production as the material foundation of society. Here is a famous passage from Marx's *Preface to A Contribution to the Critique of Political Economy* (1859):

In the social production of their existence, men inevitably enter into definite relations, which are independent of their will, namely relations of production appropriate to a given stage in the development of their material forces of production. The totality of these relations of production constitutes the economic structure of society, the real foundation, on which arises a legal and political superstructure and to which correspond definite forms of social consciousness. The mode of production of material life conditions the general process of social, political and intellectual life. It is not the consciousness of men that determines their existence, but their social existence that determines their consciousness. At a certain stage of development, the material productive forces of society come into conflict with the existing relations of production or – this merely expresses the same thing in legal terms – with the property relations within the framework of which they have operated hitherto. From forms of development of the productive forces these relations turn into their fetters. Then begins an era of social revolution. The changes in the economic foundation lead sooner or later to the transformation of the whole immense superstructure.

In studying such transformations it is always necessary to distinguish between the material transformation of the economic conditions of production, which can be determined with the precision of natural science, and the legal, political, religious, artistic or philosophic – in short, ideological forms in which men become conscious of this conflict and fight it out. Just as one does not judge an individual by what he thinks about himself, so one cannot judge such a period of transformation by its consciousness, but, on the contrary, this consciousness must be explained from the contradictions of material life, from the conflict existing between the social forces of production and the relations of production. No social order is ever *destroyed before all the productive forces for which it is sufficient have been developed, and new superior relations of production never replace older ones before the material conditions for their existence have matured within the framework of the old society.*

Mankind thus inevitably sets itself only such tasks as it is able to solve, since closer examination will always show that the problem itself arises only when the material conditions for its solution are already present or at least in the course of formation.

In broad outline, the Asiatic, ancient, feudal and modern bourgeois modes of production may be designated as epochs marking progress in the economic development of society. The bourgeois mode of production is the last antagonistic form of the social process of production – antagonistic not in the sense of individual antagonism but of an antagonism that emanates from the individuals' social conditions of existence – but the productive forces developing within bourgeois society create also the material conditions for a solution of this antagonism. The prehistory of human society accordingly closes with this social formation.

Here the emphasis is on the social relations of production, not just the forces of production (tools, materials, technical knowledge). And on this approach, study of the class relations of a given society is a "materialist" study -- even though class relations are abstract and intangible. And we provide a materialist analysis of a circumstance when we show how that circumstance corresponds to or emerges from certain features of the social relations of production.

So far, then, we seem to have two things going on: first, an approach to the history of ideas ("place systems of ideas into the roles they play in the social arrangements through which human populations satisfy material needs"), and an approach to the unfolding of history more generally ("attempt to understand historical developments in terms of the role they play in production and the satisfaction of needs"). So materialism is a theory about historical causation: what kinds of circumstances cause what other kinds of circumstances. And, perhaps, it is a theory about knowledge: that knowledge proceeds from analysis of material facts, not pure philosophical speculation or imagination.

There seem to be several hazards built into this approach. One is the temptation of reductionism that the approach seems to invite: the impulse to reduce thought, theory, and philosophy to some compound of the "needs of the social system of production". But is it really compelling to imagine that the Young Hegelians were simply working out some of the contradictions of the system of property and factory manufacturing? No; Marx's rhetoric seems to be getting away from him here -- in ways in which later thinkers such as Mannheim perhaps allowed the sociology of knowledge to spin out of control as well. And the hazard of reductionism also raises the worry of a blindspot when it comes to the relative autonomy of politics or culture: human beings seem to be better at imagining and extending political or cultural inventions than a crude materialism would permit.

So a defensible contemporary materialism can't be as simple as this: "Material conditions determine the content of culture, politics, and thought." Rather, we might hold more modestly: "Material conditions constrain, influence, and stimulate the content of culture, politics, and thought." We can understand Aristotle's philosophy better when we understand something of its material and historical setting; but the fact remains that Aristotle was a creative and imaginative philosopher who transcended his time in a variety of ways.

So, once again, what sort of theory is materialism? Perhaps we could say this: it is a "meta" - framework, a philosophical premise about how the world works. In this respect it functions as a substantive metaphysical theory. And it is a premise about how a style of thinking, a recommendation about how we should reason about the world and what factors to subject to careful analysis. Here materialism serves as something like an applied epistemology -- a theory about how and what to investigate in order to arrive at valuable, justified knowledge. It falls in the general category of ideas such as idealism, monism, atomism, physicalism, or dualism: organizing ideas about the nature of reality, within the context of which more specific theorizing and investigating can take place.

There are many questions that remain. Do these two aspects of materialism hang together? Could one accept the metaphysics but reject the epistemology, or vice versa? Once we have rejected the reductionism associated with vulgar materialism, how much remains of the theory? And is there a continuing role for materialist thinking in the twenty-first century world?

Section 7

MARC BLOCH'S MATERIALISM

One of the historians whom I most admire is Marc Bloch. He was one of France's most important medieval historians in the first half of the twentieth century, and he died at the hands of the Gestapo while serving in the Resistance in Paris in 1944. (Carole Fink's biography is an outstanding treatment of his thought and life; *Marc Bloch: A Life in History*; also important is *Marc Bloch, l'historien et la cite*.)

Here I am primarily interested in the substantive contributions Bloch brought to the writing of history. Bloch was one of the founders of the *Annales* school of history, along with Lucien Febvre, and he left a deep impression on subsequent historical imagination later in the twentieth century. In particular, he gave a strong impetus to social and sociological history, and he brought a non-Marxist materialism into the writing of history that represented a very important angle of view. The largest impact of the *Annales* school -- Febvre, Bloch, Ladurie, Braudel, Le Goff -- is the set of perspectives it forged for the understanding of social and cultural history -- looking at the structures and experiences of ordinary people as one foundation for the formation of history. This required the invention of new historical vocabulary and new sources of data. And Bloch was central in each area.

A couple of Bloch's books are most significant. *Feudal Society* is a very important contribution to our understanding of the institutions and social relations of feudalism -- the manorial system, vassalage, and kingship. And his writings about French agricultural history are of special interest (*French Rural History: An Essay on Its Basic Characteristics*). These books document quite a few important aspects of French rural social life -- both high and low. But even more importantly, Bloch brought several distinctive ideas into historical writing that continue to serve as illuminating models about how to understand the past. One is a version of materialist historical investigation -- Bloch provides great insight into the forces and relations of production in rural medieval France and the material culture of the middle ages. A second is an adept ability to single out and scrutinize some of the forms of political structure and power that defined French feudal society. And a third is a subtle way of characterizing the social whole of medieval society and mentality that owed much to Durkheim. In a curious way, then, Bloch's work picked up some of the themes that constituted modern social theory in Marx, Weber, and Durkheim.

Bloch's materialism is most evident in *French Rural History*. Here Bloch gives a detailed and scholarly treatment of the social and community consequences of the diffusion of the heavy wheeled plough. He provides a careful technical analysis of the advantages and exigencies of the heavy plough, which was most suited to the heavy soil of northern France. And he works out the social prerequisites of this technology -- basically, a degree of community organization that could successfully coordinate land use consistent with ownership and the turning radius of the heavy implement and its team of horses. The technical requirements of the plough required certain social arrangements. And the social structure of the northern French village satisfied these conditions -- in striking contrast to the looser coordination found in southern French villages. "Only a society of great compactness, composed of men who thought instinctively in terms of the community, could have created such a regime. The land itself was the fruit of collective labour" (*French Rural History*, 45).

This is materialism; but it is not especially Marxist materialism. It doesn't give primacy to class relations. And it doesn't support any kind of teleology in historical development. But the central point was clear. Bloch sought to demonstrate that a major technology -- for example, cultivation with the heavy plough -- incorporates and implicates a whole complex social and cultural system. And a major part of social history is to discover the sequence of adjustments through which the technology system is incorporated.

The Durkheim part of the story is also an important one. Durkheim was a major influence on French social thought in the teens and twenties, and the vector to Bloch was particularly direct. The journal *Annales d'Histoire Economique et Sociale* was created by Bloch and Febvre as a vehicle for inviting a more sociological approach to economic history and to encourage interdisciplinary research in this field, and Bloch and Febvre were deeply influenced by the debate that surrounded history and Durkheimian sociology in the period 1890-1910. R. Colbert Rhodes has written a good essay on Durkheim's influence on Bloch (REF). Rhodes writes: "Bloch's essentially sociological approach to historical writing is responsible for some of the most distinctive and useful features of his work. Bloch reflects the Durkheimian social realist metaphysic by reaching behind individuals to the social group considered in its broadest aspect, the collective mentality. Bloch acknowledges in the Historian's Craft his dominant interest in the study of man integrated into the social group. In the Craft, Bloch borrows a citation from Lucien Febvre to state his own interest as 'not man, again, never man.' We are interested in "human societies, organized groups" (47).

The final feature of Bloch's thought I want to highlight is his vocabulary of structure and power in his treatment of French feudalism. There is a parallel with Weber in this body of thinking. Bloch spent a year studying in Germany and was presumably aware of Weber's thought, but there is no clear evidence of direct influence. But there are several ways in which some of Bloch's thought parallels Weber's. One is in his use of ideas about historical concepts that are similar to Weber's concept of ideal types. And the other is his careful analysis of the historical realities of relations of power and social structures that embody power.

Bloch's writings repay a careful reading -- both for their importance as first-rate historical scholarship and for the light they shed on the problem of historical knowledge and conceptualization. And it is highly relevant to find that all the strands of classical sociological theory find a counterpart in his thought.

Section 8

PERSPECTIVES ON TRANSPORTATION HISTORY

April 27, 2017

https://understandingsociety.blogspot.com/2017/04/perspectives-on-transportation-history.html

I view transport as a crucial structuring condition in society that is perhaps under-appreciated and under-studied. The extension of the Red Line from Harvard Square (its terminus when I was a graduate student) to Davis Square in Somerville a decade later illustrated the transformative power of a change in the availability of urban transportation; residential patterns, the creation of new businesses, and the transformation of the housing market all shifted rapidly once it was possible to get from Davis Square to downtown Boston for a few dollars and 30 minutes. The creation of networks of super-high-speed trains in Europe and Asia does the same for the context of continental-scale economic and cultural impacts. And the advent of container shipping in the 1950's permitted a substantial surge in the globalization of the economy by reducing the cost of delivery of products from producer to consumer. Containers were a disruptive technology. It is clear that transportation systems are a crucial part of the economic, political, and cultural history of a place larger than a village; and this is true at a full range of scales.

We can look at the history of transport from several perspectives. First, we can focus on the social imperatives (including cultural values) that have influence on the development and elaboration of a transportation system. (Frank Dobbin considers some of these factors in his *Forging Industrial Policy: The United States, Britain, and France in the Railway Age*, where he considers the substantial impact that differences in political culture had on the build-out of rail networks in France, Germany, and the United States; link.) Second, we can focus on the social and political consequences that flow from the development of a new transportation system. For example, ideas and diseases spread further and faster; new population centers arise; businesses develop closer relationships with each other over greater distances. And third, we can consider the historiography of the history of transport -- the underlying assumptions that have been made by various historians who have treated transport as an important historical phenomenon.

Over fifty years ago L. Girard treated these kinds of historical effects in his contribution to *Cambridge Economic History of Europe: Volume VI (Parts I and II)*, Part I, in a chapter dedicated to "Transport". He provides attention to the main modalities of transport -- roads, sea, rail. In each case the technology and infrastructure are developed in ways that illustrate significant contingency. Consider his treatment of the development of the English road network.

> *Eventually the English network, the spontaneous product of local decisions, progressed out of this state of disorganization. Its isolated segments were linked up and ultimately provided a remarkably comprehensive network corresponding to basic national requirements. By trial and error and by comparing their processes, the trustees and their surveyors arrived at a general notion of what a road should be. (217)*

(Notice the parallels that exist between this description and the process through which the Internet was built out in the 1980s and 1990s.)

Similar comments are offered about the American rail system.

The American railroad was the product of improvisation, in contrast to the English track, which was built with great care. At first all that was required was a fairly rough and ready line which could operate with a minimum amount of equipment. Then as traffic increased and profits began to be made, the whole enterprise was transformed to take account of the requirements of increased traffic and of the greater financial possibilities. (232)

Despite all their improvisations and wastage, the American railroads astonished Europe, which saw a whole continent come to life in the path of the lines. The railways opened up America for a second time. By 1850 the east-west link between western Europe and the Mississippi valley was already created by means of the States on the Atlantic seaboard. The supremacy of the Chicago-New York axis had become established, at the expense of the South and of Canada, which were taking more time to get organized. America swung away from a north-south to an east-west orientation. (233)

And here is a somewhat astounding claim:

The northern railways allowed the Union to triumph in the Civil War, which was fought in part to determine the general direction to be taken by the future railways. (233-234)

Also surprising is the role that Girard attributes to the politics of railroads in the ascendancy of Napoleon III in 1851 (239).

Here is Girard's summary of the large contours of the development of transport during this critical period:

Whatever the course of future history, the century of the railway and the steamer marks a decisive period in the history of transport, and that of the world. Particular events in political history often tend to assume less and less importance as time goes on. But the prophecies of Saint-Simon on the unification of the planet, and the meeting of the races for better or for *worse, remain excitingly topical. Man has changed the world, and the world has changed man -- in a very short time indeed. (273)*

This history was written in 1965, over fifty years ago. One thing that strikes the contemporary reader is how disinterested the author appears to be in cause and effect. He does not devote much effort to the question, what forces drove the discoveries and investments that resulted in a world-wide network of railways and steamships? And he does not consider in any substantial detail the effects of this massive transformation of activities at a national and global scale. Further, Gerard gives no indication of interest in the social context or setting of transport -- how transport interacted with ordinary people, how it altered the environment of everyday life, how it contributed to social problems and social solutions. It seems reasonable to believe that the history of transport during this period would be written very differently today.

Chapter 4

ETHICS AND TECHNOLOGY

AN EXISTENTIAL PHILOSOPHY OF TECHNOLOGY

Ours is a technological culture, at least in the quarter of the countries in the world that enjoy a high degree of economic affluence. Cell phones, computers, autonomous vehicles, CT scan machines, communications satellites, nuclear power reactors, artificial DNA, artificial intelligence bots, drone swarms, fiber optic data networks -- we live in an environment that depends unavoidably upon complex, scientifically advanced, and mostly reliable artifacts that go well beyond the comprehension of most consumers and citizens. We often do not understand how they work. But more than that, we do not understand how they affect us in our social, personal, and philosophical lives. We are different kinds of persons than those who came before us, it often seems, because of the sea of technological capabilities in which we swim. We think about our lives differently, and we relate to the social world around us differently.

How can we begin investigating the question of how technology affects the conduct of a "good life"? Is there such a thing as an "existential" philosophy of technology -- that is, having to do with the meaning of the lives of human beings in the concrete historical and technological circumstances in which we now find ourselves? This suggests that we need to consider a particularly deep question: in what ways does advanced technology facilitate the good human life, and in what ways does it frustrate and block the good human life? Does advanced technology facilitate and encourage the development of full human beings, and lives that are lived well, or does it interfere with these outcomes?

We are immediately drawn to a familiar philosophical question, What is a good life, lived well? This has been a central question for philosophers since Aristotle and Epicurus, Kant and Kierkegaard, Sartre and Camus. But let's try to answer it in a paragraph. Let's postulate that there are a handful of characteristics that are associated with a genuinely valuable human life. These might include the individual's realization of a capacity for self-rule, creativity, compassion for others, reflectiveness, and an ability to grow and develop. This suggests that we start from the conception of a full life of freedom and development offered by Amartya Sen in Development as Freedom and the list of capabilities offered by Martha Nussbaum in Creating Capabilities: The Human Development Approach -- capacities for life, health, imagination, emotions, practical reason, affiliation with others, and self-respect. And we might say that a "life lived well" is one in which the person has lived with integrity, justice, and compassion in developing and fulfilling his or her fundamental capacities. Finally, we might say that a society that enables the development of each of these capabilities in all its citizens is a good society.

Now look at the other end of the issue -- what are some of the enhancements to human living that are enabled by modern technologies? There are several obvious candidates. One might say that technology facilitates learning and the acquisition of knowledge; technology can facilitate health (by finding cures and preventions of disease; and by enhancing nutrition, shelter, and other necessities of daily life); technology can facilitate human interaction (through the forms of communication and transportation enabled by modern technology); technology can enhance compassion by acquainting us with the vivid life experiences of others. So technology is sometimes life-enhancing and fulfilling of some of our most fundamental needs and capabilities.

How might Dostoevsky, Dos Passos, Baldwin, or Whitman have adjusted their life plans if confronted by our technological culture? We would hope they would not have been overwhelmed in their imagination and passion for discovering the human in the ordinary by an iPhone, a Twitter feed, and a web browser. We would like to suppose that their insights and talents would have survived and flourished, that poetry, philosophy, and literature would still have emerged, and that compassion and commitment would have found its place even in this alternative world.

But the negative side of technology for human wellbeing is also easy to find. We might say that technology encourages excessive materialism; it draws us away from real interactions with other human beings; it promotes a life consisting of a series of entertaining moments rather than meaningful interactions; and it squelches independence, creativity, and moral focus. So the omnipresence of technologies does not ensure that human beings will live well and fully, by the standards of Aristotle, Epicurus, or Montaigne.

In fact, there is a particularly bleak possibility concerning the lives that advanced everyday technology perhaps encourages: our technological culture encourages us to pursue lives that are primarily oriented towards material satisfaction, entertainment, and toys. This sounds a bit like a form of addiction or substance abuse. We might say that the ambient cultural imperatives of acquiring the latest iPhone, the fastest internet streaming connection, or a Tesla are created by the technological culture that we inhabit, and that these motivations are ultimately unworthy of a fully developed human life. Lucretius, Socrates, and Montaigne would scoff.

It is clear that technology has the power to distort our motives, goals and values. But perhaps with equal justice one might say that this is a life world created by capitalism rather than technology -- a culture that encourages and elicits personal motivations that are "consumerist" and ultimately empty of real human value, a culture that depersonalizes social ties and trivializes human relationships based on trust, loyalty, love, or compassion. This is indeed the critique offered by theorists of the philosophers of the Frankfurt School -- that capitalism depends upon a life world of crass materialism and impoverished social and personal values. And we can say with some exactness how capitalism distorts humanity and culture in its own image: through the machinations of advertising, strategic corporate communications, and the honoring of acquisitiveness and material wealth (link). It is good business to create an environment where people want more and more of the gadgets that technological capitalism can provide.

So what is a solution for people who worry about the shallowness and vapidity of this kind of technological materialism? We might say that an antidote to excessive materialism and technology fetishism is a fairly simple maxim that each person can strive to embrace: aim to identify and pursue the things that genuinely matter in life, not the glittering objects of short-term entertainment and satisfaction. Be temperate, reflective, and purposive in one's life pursuits. Decide what values are of the greatest importance, and make use of technology to further those values, rather than as an end in itself. Let technology be a tool for creativity and commitment, not an end in itself. Be selective and deliberate in one's use of technology, rather than being the hapless consumer of the latest and shiniest. Create a life that matters.

Section 1

ETHICAL PRINCIPLES FOR ASSESSING NEW TECHNOLOGIES

Technologies and technology systems have deep and pervasive effects on the human beings who live within their reach. How do normative principles and principles of social and political justice apply to technology? Is there such a thing as "the ethics of technology"?

There is a reasonably active literature on questions that sound a lot like these. (See, for example, the contributions included in Winston and Edelbach, eds., *Society, Ethics, and Technology*.) But all too often the focus narrows too quickly to ethical issues raised by a particular example of contemporary technology -- genetic engineering, human cloning, encryption, surveillance, and privacy, artificial intelligence, autonomous vehicles, and so forth. These are important questions; but it is also possible to ask more general questions as well, about the normative space within which technology, private activity, government action, and the public live together. What principles allow us to judge the overall justice, fairness, and legitimacy of a given technology or technology system?

Consider three large principles that have emerged in other areas of social and political ethics as a basis for judging the legitimacy and fairness of a given set of social arrangements:

A. Technologies should contribute to some form of human good, some activity or outcome that is desired by human beings -- health, education, enjoyment, pleasure, sociality, friendship, fitness, spirituality, ...

B. Technologies ought to be consistent with the fullest development of the human capabilities and freedoms of the individuals whom they affect. [Or stronger: “promote the fullest development …”]

C. Technologies ought to have population effects that are fair, equal, and just.

The first principle attempts to address the question, "What is technology good for? What is the substantive moral good that is served by technology development?" The basic idea is that human beings have wants and needs, and contributing to their ability to fulfill these wants is itself a good thing (if in so doing other greater harms are not created as well). This principle captures what is right about utilitarianism and hedonism -- the inherent value of human happiness and satisfaction. This means that entertainment and enjoyment are legitimate goals of technology development.

The second principle links technology to the "highest good" of human wellbeing -- the full development of human capabilities and freedoms. As is evident, the principle offered here derives from Amartya Sen's theory of capabilities and functionings, expressed in *Development as Freedom*. This principle recalls Mill's distinction between higher and lower pleasures:

Mill always insisted that the ultimate test of his own doctrine was utility, but for him the idea of the greatest happiness of the greatest number included qualitative judgements about different levels or kinds of human happiness. Pushpin was not as good as poetry; only Pushkin was.... Cultivation of one's own individuality should be the goal of human existence. (J.S. McClelland, A History of Western Political Thought : 454)

The third principle addresses the question of fairness and equity. Thinking about justice has evolved a great deal in the past fifty years, and one thing that emerges clearly is the intimate connection between injustice and invidious discrimination -- even if unintended. Social institutions that arbitrarily assign significantly different opportunities and life outcomes to individuals based on characteristics such as race, gender, income, neighborhood, or religion are unfair and unjust, and need to be reformed. This approach derives as much from current discussions of racial health disparities as it does from philosophical theories along the lines of Rawls and Sen.

On these principles a given technology can be criticized, first, if it has no positive contribution to make for the things that make people happy or satisfied; second, if it has the effect of stunting the development of human capabilities and freedoms;

and third, if it has discriminatory effects on quality of life across the population it effects.

One important puzzle facing the ethics of technology is a question about the intended audience of such a discussion. We are compelled to ask, to whom is a philosophical discussion of the normative principles that ought to govern our thinking about technology aimed? Whose choices, actions, and norms are we attempting to influence? There appear to be several possible answers to this question.

Corporate ethics. Entrepreneurs and corporate boards and executives have an ethical responsibility to consider the impact of the technologies that they introduce into the market. If we believe that codes of corporate ethics have any real effect on corporate decision-making, then we need to have a basis in normative philosophy for a relevant set of principles that should guide business decision-making about the creation and implementation of new technologies by businesses. A current example is the use of facial recognition for the purpose of marketing or store security; does a company have a moral obligation to consider the negative social effects it may be promoting by adopting such a technology?

Governments and regulators. Government has an overriding responsibility of preserving and enhancing the public good and minimizing harmful effects of private activities. This is the fundamental justification for government regulation of industry. Since various technologies have the potential of creating harms for some segments of the public, it is legitimate for government to enact regulatory systems to prevent reckless or unreasonable levels of risk. Government also has a responsibility for ensuring a fair and just environment for all citizens, and enacting policies that serve to eliminate inequalities based on discriminatory social institutions. So here too governments have a role in regulating technologies, and a careful study of the normative principles that should govern our thinking about the fairness and justice of technologies is relevant to this process of government decision-making as well.

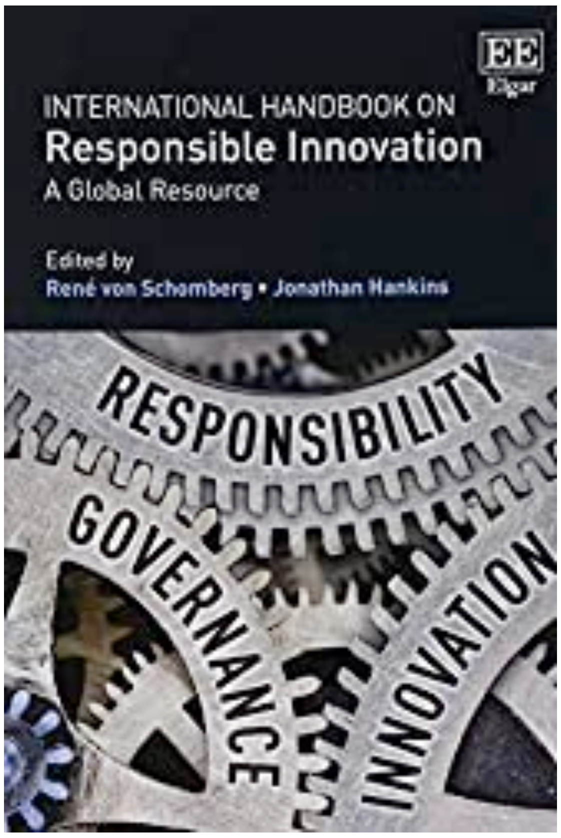
Public interest advocacy groups. One way in which important social issues can be debated and sometimes resolved is through the advocacy of well-organized advocacy groups such as the Union of Concerned Scientists, the Sierra Club, or Greenpeace. Organizations like these are in a position to argue in favor of or against a variety of social changes, and raising concerns about specific kinds of technologies certainly falls within this scope. There are only a small number of grounds for this kind of advocacy: the innovation will harm the public, the innovation will create unacceptable hidden costs, or the innovation raises unacceptable risks of unjust treatment of various groups. In order to make the latter kind of argument, the advocacy group needs to be able to articulate a clear and justified argument for its position about "unjust treatment".

The public. Citizens themselves have an interest in being able to make normative judgments about new technologies as they arise. "This technology looks as though it will improve life for everyone and should be favored; that technology looks as though it will create invidious and discriminatory sets of winners and losers and should be carefully regulated." But for citizens to have a basis for making judgments like these, they need to have a normative framework within which to think and reason about the social role of technology. Public discussion of the ethical principles underlying the legitimacy and justice of technology innovations will deepen and refine these normative frameworks.

Considered as proposed here, the topic of "ethics of technology" is part of a broad theory of social and political philosophy more generally. It invokes some of our best reasoning about what constitutes the human good (fulfillment of capabilities and freedoms) and about what constitutes a fair social system (elimination of invidious discrimination in the effects of social institutions on segments of population). Only when we have settled these foundational questions are we able to turn to the more specific issues often discussed under the rubric of the ethics of technology.

Section 2

RESPONSIBLE INNOVATION



January 26, 2020

https://understandingsociety.blogspot.com/2020/01/responsible-innovation-and-philosophy.html

Several posts here have focused on the philosophy of technology (link, link, link, link). A simple definition of the philosophy of technology might go along these lines:

Technology may be defined broadly as the sum of a set of tools, machines, and practical skills available at a given time in a given culture through which human needs and interests are satisfied and the interplay of power and conflict furthered. The philosophy of technology offers an interdisciplinary approach to better understanding the role of technology in society and human life. The field raises critical questions about the ways that technology intertwines with human life and the workings of society. Do human beings control technology? For whose benefit? What role does technology play in human wellbeing and freedom? What role does technology play in the exercise of power? Can we control technology? What issues of ethics and social justice are raised by various technologies? How can citizens within a democracy best ensure that the technologies we choose will lead to better human outcomes and expanded capacities in the future?

One of the issues that arises in this field is the question of whether there are ethical principles that should govern the development and implementation of new technologies. (This issue is discussed further in an earlier post; link.)

One principle of technology ethics seems clear: policies and regulations are needed to protect the future health and safety of the public. This is the same principle that serves as the ethical basis of government regulation of current activities, justifying coercive rules that prevent pollution, toxic effects, fires, radiation exposure, and other clear harms affecting the health and safety of the public.

Another principle might be understood as exhortatory rather than compulsory, and that is the general recommendation that technologies should be pursued by private actors that make some positive contribution to human welfare. This principle is plainly less universal and obligatory than the "avoid harm" principle; many technologies are chosen because their inventors believe they will entertain, amuse, or otherwise please members of the public, and will thereby permit generation of profits. (Here is a discussion of the value of entertainment; link.)

A more nuanced exhortation is the idea that inventors and companies should subject their technology and product innovation research to broad principles of sustainability. Given that large technological change can potentially have very large environmental and collective effects, we might think that companies and inventors should pay attention to the large challenges our society faces, now and in the foreseeable future: addiction, obesity, CO_2 production, plastic waste, erosion of privacy, spread of racist politics, fresh water depletion, and information disparities, to name several.

These principles fall within the general zone of the ethics of corporate social responsibility. Many companies pay lip service to the social-benefits principle and the sustainability principle, though it is difficult to see evidence of the effectiveness of this motivation. Business interests often seem to trump concerns for positive social effects and sustainability -- for example, in the pharmaceutical industry and its involvement in the opioid crisis (link).

It is in the context of these reflections about the ethics of technology that I was interested to learn of an academic and policy field in Europe called "responsible innovation". This is a network of academics, government officials, foundations, and non-profit organizations working together to try to induce more directionality in technology change (innovation). René von Schomberg and Jonathan Hankins's recently published volume *International Handbook on Responsible Innovation: A Global Resource* gives an in-depth exposure to the thinking, research, and policy advocacy that this network has accumulated. A key actor in the advancement of this field has been the Bassetti Foundation (link) in Milan, which has made the topic of responsible innovation central to its mission for several decades. *The Journal of Responsible Innovation* provides a look at continuing research in this field.

The primary locus of discussion and applications in the field of RRI has been within the EU. There is not much evidence of involvement in the field from United States actors in this movement, though the Virtual Institute of Responsible Innovation at Arizona State University has received support from the US National Science Foundation (link).

Von Schomberg describes the scope and purpose of the RRI field in these terms:

Responsible Research and Innovation is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society). (2)

The definition of this field overlaps quite a bit with the philosophy and ethics of technology, but it is not synonymous. For one thing, the explicit goal of RRI is to help provide direction to the social, governmental, and business processes driving innovation. And for another, the idea of innovation isn't exactly the same as "technology change". There are social and business innovations that fall within the scope of the effort — for example, new forms of corporate management or new kinds of financial instruments -- but which do not fall within the domain of technological innovations.

Von Schomberg has been a leading thinker within this field, and his contributions have helped to set the agenda for the movement. In his contribution to the volume he identifies six deficits in current innovation policy in Europe (all drawn from chapter two of the volume):

1. Exclusive focus on risk and safety issues concerning new technologies under governmental regulations
2. Market deficits in delivering on societal desirable innovations
3. Aligning innovations with broadly shared public values and expectations
4. A focus on the responsible development of technology and technological potentials rather than on responsible innovations
5. A lack of open research systems and open scholarship as a necessary, but not sufficient condition for responsible innovation
6. Lack of foresight and anticipative governance for the alternative shaping of innovation in sectors

Each of these statements involves very complex ideas about society-government-corporate relationships, and we may well come to judge that some of the recommendations made by Schomberg are more convincing than others. But the clarity of this statement of the priorities and concerns of the RRI movement is enormously valuable as a way of advancing debate on the issues.

The examples that von Schomberg and other contributors discuss largely have to do with large innovations that have sparked significant public discussion and opposition — nuclear power, GMO foods, nanotechnology-based products. These example focus attention on the later stages of scientific and technological knowledge when it comes to the point of introducing the technology into the public. But much technological innovation takes place at a much more mundane level -- consumer electronics and software, enhancements of solar technology, improvements in electric vehicle technology, and digital personal assistants (Alexa, Siri), to name a few.

A defining feature of the RRI field is the explicit view that innovation is not inherently good or desirable (for example, in the contribution by Luc Soete in the volume). Contrary to the assumptions of many government economic policy experts, the RRI network is unified in criticism of the idea that innovation is always or usually productive of economic growth and employment growth. These observers argue instead that the public should have a role in deciding which technological options ought to be pursued, and which should not.

In reading the programmatic statements of purpose offered in the volume, it sometimes seems that there is a tendency to exaggerate the degree to which scientific and technological innovation is (or should be) a directed and collectively controlled process. The movement seems to undervalue the important role that creativity and invention play within the crucial fact of human freedom and fulfillment. It is an important moral fact that individuals have extensive liberties concerning the ways in which they use their talents, and the presumption needs to be in favor of their right to do so without coercive interference. Much of what goes on in the search for new ideas, processes, and products falls properly on the side of liberty rather than a socially regulated activity, and the proper relation of social policy to these activities seems to be one of respect for the human freedom and creativity of the innovator rather than a prescriptive and controlling one. (Of course some regulation and oversight is needed, based on assessments of risk and harm; but von Schomberg and others dismiss this moral principle as too limited.)

It sometimes seems as though the contributors slide too quickly from the field of government-funded research and development (where the public has a plain interest in "directing" the research at some level), to the whole ecology of innovation and discovery, whether public, corporate, or academic. As noted above, von Schomberg considers the governmental focus on harm and safety to be the "first deficit" — in other words, an insufficient basis for "guiding innovation". In contrast, he wants to see public mechanisms tasked with "redirecting" technology innovations and industries. However, much innovation is the result of private initiative and funding, and it seems that this field appropriately falls outside of prescription by government (beyond normal harm-based regulatory oversight). Von Schomberg uses the phrase "a proper embedding of scientific and technological advances in society"; but this seems to be a worrisome overreach, in that it seems to imply that all scientific and technology research should be guided and curated by a collective political process.

This suggests that a more specific description of the goals of the movement would be helpful. Here is one possible specification:

- Require government agencies to justify the funding and incentives that they offer in support of technology innovation based on an informed assessment of the public's preferences;
- Urge corporations to adopt standards to govern their own internal innovation investments to conform to acknowledged public concerns (environmental sustainability, positive contributions to health and safety of citizens and consumers, ...);
- Urge scientists and researchers to engage in public discussion of their priorities in scientific and technological research.
- Create venues for open and public discussion of major technological choices facing society in the current century, leading to more articulate understanding of priorities and risks.

There is an interesting parallel here with the Japanese government's efforts in the 1980s to guide investment and research and development resources into the highest priority fields to advance the Japanese economy. The US National Research Council study, 21st Century Innovation Systems for Japan and the United States: Lessons from a Decade of Change: Report of a Symposium (2009) (link), provides an excellent review of the strategies adopted by the United States and Japan in their efforts to stimulate technology innovation in chip production and high-end computers from the 1960s to the 1990s. These efforts were entirely guided by the effort to maintain commercial and economic advantage in the global marketplace. Jason Owen-Smith addresses the question of the role of US research universities as sites of technological research in *Research Universities and the Public Good: Discovery for an Uncertain Future*; link.

The "responsible research and innovation" (RRI) movement in Europe is a robust effort to pose the question, how can public values be infused into the processes of technology innovation that have such a massive potential effect on public welfare? It would seem that a major aim of the RRI network is to help to inform and motivate commitments by corporations to principles of responsible innovation within their definitions of corporate social responsibility, which is unmistakably needed. It is worthwhile for U.S. policy experts and technology ethicists alike to pay attention to these debates in Europe, and the *International Handbook on Responsible Innovation* is an excellent place to begin.

Section 3

ETHICAL DISASTERS

Many examples of technical disasters have been provided in *Understanding Society*, along with efforts to understand the systemic dysfunctions that contributed to their occurrence. Frequently those dysfunctions fall within the business organizations that manage large, complex technology systems, and often enough those dysfunctions derive from the imperatives of profit-maximization and cost avoidance. Andrew Hopkins' account of the business decisions contributing to the explosion of the ESSO gas plant in Longford, Australia illustrates this dynamic in *Lessons from Longford: The ESSO Gas Plant Explosion*. The withdrawal of engineering experts from the plant to a remote corporate headquarters was a cost-saving move that, according to Hopkins, contributed to the eventual disaster.

A topic we have not addressed in detail is the occurrence of ethical disasters -- terrible outcomes that are the result of deliberate choices by decision-makers within an organization that are, upon inspection, clearly and profoundly unethical and immoral. The collapse of Enron is probably one such disaster; the Bernie Madoff scandal is another. But it seems increasingly likely that Purdue Pharma and the Sackler family's business leadership of the corporation represent another major example. Recent reporting by *ProPublica*, the *Atlantic*, and the *New York Times* relies on documents collected in the course of litigation against Purdue Pharma and members of the Sackler family in Massachusetts and New York. (Here are the unredacted court documents on which much of this reporting depends [REF].) These documents make it hard to avoid the ethical conclusion that the Sackler family actively participated in business strategies for their company Purdue Pharma that treated the OxyContin addiction epidemic as an expanding business opportunity. And this seems to be a huge ethical breach.

This set of issues is currently unresolved by the courts, so it rests with the legal system to resolve the facts and the issues of legal culpability. But as citizens we all have the ability to read the documents and make our own decisions about the ethical status of decisions and strategies made by the family and the corporation over the course of this disaster. The point here is simply to ask these key questions: how should we think about the ethical status of decisions and strategies of owners and managers that lead to terrible harms, and harms that could reasonably have been anticipated? How should a company or a set of owners respond to a catastrophe in which several hundred thousand people have died, and which was facilitated in part by deliberate marketing efforts by the company and the owners? How should the company have adjusted its business when it became apparent that its product was creating addiction and widespread death?

First, here are a few details from the current reporting about the case. Here are a few paragraphs from the *ProPublica* story (January 30, 2019):

Not content with billions of dollars in profits from the potent painkiller OxyContin, its maker explored expanding into an "attractive market" fueled by the drug's popularity — treatment of opioid addiction, according to previously secret passages in a court document filed by the state of Massachusetts.

In internal correspondence beginning in 2014, Purdue Pharma executives discussed how the sale of opioids and the treatment of opioid addiction are "naturally linked" and that the company should expand across "the pain and addiction spectrum," according to redacted sections of the lawsuit by the Massachusetts attorney general. A member of the billionaire Sackler family, which founded and controls the privately held company, joined in those discussions and urged staff in an email to give "immediate attention" to this business opportunity, the complaint alleges. (ProPublica 1/30/2019; link)

The NYT story reproduces a diagram included in the New York court filings that illustrates the company's business strategy of "Project Tango" -- the idea that the company could make money both from sales of its pain medication and from sales of treatments for the addiction it caused.

Further, according to the reporting provided by the NYT and ProPublica, members of the Sackler family used their positions on the Purdue Pharma board to press for more aggressive business exploitation of the opportunities described here:

In 2009, two years after the federal guilty plea, Mortimer D.A. Sackler, a board member, demanded to know why the company wasn't selling more opioids, email traffic cited by Massachusetts prosecutors showed. In 2011, as states looked for ways to curb opioid prescriptions, family members peppered the sales staff with questions about how to expand the market for the drugs.... The family's statement said they were just acting as responsible board members, raising questions about "business issues that were highly relevant to doctors and patients. (NYT 4/1/2019; link)

From the 1/30/2019 *ProPublica* story, and based on more court documents:

Citing extensive emails and internal company documents, the redacted sections allege that Purdue and the Sackler family went to extreme lengths to boost OxyContin sales and burnish the drug's reputation in the face of increased regulation and growing public awareness *of its addictive nature. Concerns about doctors improperly prescribing the drug, and patients becoming addicted, were swept aside in an aggressive effort to drive OxyContin sales ever higher, the complaint alleges. (link)*

And *ProPublica* underlines the fact that prosecutors believe that family members have personal responsibility for the management of the corporation:

The redacted paragraphs leave little doubt about the dominant role of the Sackler family in Purdue's management. The five Purdue directors who are not Sacklers always voted with the family, according to the complaint. The family-controlled board approves everything from the number of sales staff to be hired to details of their bonus incentives, which have been tied to sales volume, the complaint says. In May 2017, when longtime employee Craig Landau was seeking to become Purdue's chief executive, he wrote that the board acted as "de-facto CEO." He was named CEO a few weeks later. (link)

The courts will resolve the question of legal culpability. The question here is one of the ethical standards that should govern the actions and strategies of owners and managers. Here are several simple ethical observations that seem relevant to this case.

First, it is obvious that pain medication is a good thing when used appropriately under the supervision of expert and well-informed physicians. Pain management enhances quality of life for people experiencing pain.

Second, addiction is plainly a bad thing, and it is worse when it leads to predictable death or disability for its victims. A company has a duty of concern for the quality of life of human beings affected by its product, and this extends to a duty to take all possible precautions to minimize the likelihood that human beings will be harmed by the product.

Third, given that the risks of addiction that were known about this product, the company has a moral obligation to treat its relations with physicians and other health providers as occasions of accurate and truthful education about the product, not opportunities for persuasion, inducement, and marketing. Rather than a sales force of representatives whose incomes are determined by the quantity of the product they sell, the company has a moral obligation to train and incentivize its representatives to function as honest educators providing full information about the risks as well as the benefits of the product. And, of course, it has an obligation not to immerse itself in the dynamics of "conflict of interest" discussed elsewhere (REF) -- this means there should be no incentives provided to the physicians who agree to prescribe the product.

Fourth, it might be argued that the profits generated by the business of a given pharmaceutical product should be used proportionally to ameliorate the unavoidable harms it creates. Rather than making billions in profits from the sale of the product, and then additional hundreds of millions on products that offset the addictions and illness created by dissemination of the product (this was the plan advanced as "Project Tango"), the company and its owners should hold themselves accountable for the harms created by their product. (That is, the social and human costs of addiction should not be treated as "externalities" or even additional sources of profit for the company.)

Finally, there is an important question at a more individual scale. How should we think about super-rich owners of a company who seem to lose sight entirely of the human tragedies created by their company's product and simply demand more profits, more timely distribution of the profits, and more control of the management decisions of the company? These are individual human beings, and surely they have a responsibility to think rigorously about their own moral responsibilities. The documents released in these court proceedings seem to display an amazing blindness to moral responsibility on the part of some of these owners.

(Stephen Arbogast's *Resisting Corporate Corruption: Cases in Practical Ethics From Enron Through The Financial Crisis* is an interesting source on corporate ethics.)

Section 4

Conflicts of interest

The possibility or likelihood of conflict of interest is present in virtually all professions and occupations. We expect a researcher, a physician, or a legislator to perform her work according to the highest values and norms of their work (searching for objective knowledge, providing the best care possible for the patient, drafting and supporting legislation in order to enhance the public good). But there is always the possibility that the individual may have private financial interests that distort or bias the work she does, and there may be large companies that have a financial interest in one set of actions rather than another.

Marc Rodwin's *Conflicts of Interest and the Future of Medicine: The United States, France, and Japan* is a rigorous and fair treatment of this issue with respect to conflicts of interest in the field of medicine. Rodwin has published extensively on this topic, and the current book is an important exploration of how professional ethics, individual interest, and business and professional institutions intersect to influence practitioner behavior in this field. The institutional actors in this story include the pharmaceutical companies and medical device manufacturers, insurers, hospitals and physician partnerships, and legislators and regulators. Rodwin shows in detail how differences in insurance policies, physician reimbursement policies, and gifts and benefits from health-related businesses to physicians contribute to an institutional environment where the physician's choices are all too easily influenced by considerations other than the best health outcomes of the patient. Rodwin finds that the institutional setting for health economics is different in the US, France, and Japan, and that these differences lead to differences in physician behavior.

Here is Rodwin's clear statement of the material situation that creates the possibility or likelihood of conflicts of interest in medicine.

Physicians earn their living through their medical work and so may practice in ways that enhance their income rather than the interests of patients. Moreover, when physicians prescribe drugs, devices, and treatments and choose who supplies these or refer patients to other providers, they affect the fortunes of third parties. As a result, providers, suppliers, and in-

surers try to influence physicians' clinical decisions for their own benefit. Thus, at the core of doctoring lies tension between self-interest and faithful service to patients and the public. The prevailing powerful medical ethos does influence physicians. Still, there is conflict between professional ethics and financial incentives. (kl 251)

Jerome Kassirer is a former editor-in-chief of the *New England Journal of Medicine*, and an expert observer of the field, and he provided a foreword to the book. Kassirer describes the current situation in the medical economy in these terms, drawing on his own synthesis of recent research and journalism:

Professionalism had been steadily eroded by complex financial ties between practicing physicians and academic physicians on the one hand and the pharmaceutical, medical device, and biotechnology industries on the other. These financial ties were deep and wide: they threatened to bias the clinical research on which physicians relied to care for the sick, and they permeated nearly every aspect of medical care. Physicians were accepting gifts, taking free trips, serving on companies' speakers' bureaus, signing their names to articles written for them by industry-paid ghostwriters, and engaging in research that endangered patient care. (kl 73)

The fundamental problem posed by Rodwin's book is this set of questions:

In what context can physicians be trusted to act in their patients' interests? How can medical practice be organized to minimize physicians' conflicts of interest? How can society promote what is best in medical professionalism? What roles should physicians and organized medicine play in the medical economy? What roles should insurers, the state, and markets play in medical care? (kl 267)

The book sheds light on dozens of institutional arrangements that create the likelihood of conflicted choices, or that reduce that likelihood. One of those arrangements is the question for a non-profit hospital of whether the physicians are employed with a fixed salary or work on a fee-for-service basis. The latter system gives the physician a very different set of financial interests, including the possibility of making clinical choices that increase revenues to the physician or his or her group practice.

Consider physicians employed as public servants in public hospitals. Typically, they receive a fixed salary set by rank, enjoy tenure, and have clinical discretion. As a result, they lack financial incentives that bias their choices and have clinical freedom. Such arrangements preclude employment conflicts of interest. But relax some of these conditions and employers can compromise medical practice.... Furthermore, emplloyers can manage physicians to promote the organization's goals. As a result, employed physicians might practice in ways that promote their employer's over their patients' interests. (kl 445)

And the disadvantages for the patient of the self-employed physician are also important:

Payment can encourage physicians to supply more, less, or different kinds of services, or to refer to particular providers. Each form of payment has some bias, but some compromise clinical decisions more than others do. (kl 445)

Plainly, the circumstances and economic institutions described here are relevant to many other occupations as well. Scientists, policymakers, regulators, professors, and accountants all face similar circumstances -- though the financial stakes in medicine are particularly high. (Here is an earlier post on corporate efforts to influence scientific research; link.)

This field of research makes an important contribution to a particular challenging topic in contemporary healthcare. But Rodwin's study also provides an important contribution to the new institutionalism, since it serves as a micro-level case study of the differences in behavior created by differences in institutional rules and practices.

Each country's laws, insurance, and medical institutions shape medical practice; and within each country, different forms of practice affect clinical choices. (kl 218)

This feature of the book allows it to contribute to the kinds of arguments on the causal and historical importance of specific configurations of institutions offered by Kathleen Thelen (link) and Frank Dobbin (link).

Section 5

ENTERTAINMENT AS A VALUABLE THING

December 18, 2019

https://understandingsociety.blogspot.com/2019/12/entertainment-as-valuable-thing.html

Quite a bit of the GDP of the United States goes into a broad category we can call "entertainment" -- television, video streaming services, books and newspapers, concerts, theatre, sports events (live and broadcast), and video games. The entertainment industry amounts to $717 billion in the US economy (link), and professional athletics adds another $73.5 billion (link). This category approaches a trillion dollars, and we haven't even taken account of the video gaming sector. US GDP is about $19.4 trillion; so entertainment in all sectors may amount to 5% or more of the total US economy.

It is interesting to take a look at the Bureau of Labor Statistics American Time Use Survey (link) to get an empirical idea of how Americans of varying ages spend their time each day. Table 8A breaks down "time spent in primary activities for the civilian population 18 years and over". Out of a 24-hour day, personal care activities and sleeping amount to 18 hours per day; organizational, civic, and religious activities amount to .27 hours per day; and leisure and sports activities amount to 4.12 hours per day. Work takes 3.94 hours per day (on average; 4.96 hours for men and 3.07 hours for women). (These are averages over a population, which explains the relatively low number of hours spent working.)

Both these observations demonstrate that millions of people value entertainment: they pay for it and the spend their time engaging with it. So entertainment is plainly valued in the lives of most people. But we can still ask a more fundamental question: what is the good of entertainment? Should society be organized in such a way as to facilitate the ability of people to find and consume sources of entertainment? Is more entertainment better than less?

There are a few simple answers to the question.

First, people often want to be entertained and choose to be entertained, and it is a basic principle of liberalism that it is good for people to exercise their liberty by doing what they want to do. (Liberty principle)

Second, people take pleasure in being entertained, pleasure is an important component of happiness, and happiness is a good thing. Therefore entertainment is a good thing, and people should be in a position to be entertained. (Happiness principle)

Third, sometimes "entertainment" is developmental and fulfilling of human capabilities for imagination, empathy, and ethical reasoning. By watching *Citizen Kane* or reading *The Fire Next Time*, people gain new insights into their own lives and the lives of others; they extend their capacity for empathy; and they enrichen their ability to think about complex social and moral topics. Martha Nussbaum makes this point in her discussions of the value of literature in *Poetic Justice: The Literary Imagination and Public Life* and *Cultivating Humanity: A Classical Defense of Reform in Liberal Education*. Nussbaum's views are discussed by Heather McRobie in *OpenDemocracy* (link). These outcomes too seem intrinsically valuable by the individuals who experience them. This view of the potential value of "entertainment" converges with the capabilities approach to human wellbeing. (Human development principle)

But, of course, not all forms of entertainment are uplifting in an intellectual or moral sense. Six seasons of The Sopranos probably didn't result in much intellectual or moral uplift for the millions of viewers who enjoyed the series, and an earlier generation's interest in *Gunsmoke*, *Bonanza*, and *Car 54* was probably equally unrewarded. So what about "pure entertainment", without moral, intellectual, or aesthetic redeeming value? What about Tetris, Solitaire, Grand Theft Auto, or the Madden NFL franchise? What about *Downton Abbey*?

J. S. Mill, far-sighted media critic that he was, distinguished between higher and lower pleasures, and argued that the former are inherently preferable to any agent who has experienced both over the latter. His reasoning is based on his view that human beings have higher faculties and lower faculties; the higher pleasures exercise the higher faculties; and anyone who has experienced both will prefer the higher pleasures. Here are a few lines from *Utilitarianism*.

Now, it is an unquestionable fact that the way of life that employs the higher faculties is strongly preferred to the way of life that caters only to the lower ones· by people who are equally acquainted with both and equally capable of appreciating and enjoying both. Few human creatures would agree to be changed into any of the lower animals in return for a promise of the fullest allowance of animal pleasures; no intelligent human being would consent to be a fool, no educated person would prefer to be an ignoramus, no person of feeling and conscience would rather be selfish and base, even if they were convinced that the fool, the dunce or the rascal is better satisfied with his life than they are with theirs. . . . If they ever think they would, it is only in cases of unhappiness so extreme that to escape from it they would exchange their situation for almost any other, however undesirable they may think the other to be. Someone with higher faculties requires more to make him happy, is probably capable of more acute suffering, and is certainly vulnerable to suffering at more points, than someone of an inferior type; but in spite of these drawbacks he can't ever really wish to sink into what he feels to be a lower grade of existence. .... It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied. And if the fool, or the pig, are of a different opinion, it is because they only know their own side of the question. The other party to the comparison knows both sides. (14)

This position is based on Mill's idea that people who have experienced both kinds of pleasures will choose the higher (demanding, challenging, aesthetically and morally complex) pleasures over the lower pleasures. But this view seems to be largely refuted by the current entertainment industry; the audience for bad television certainly includes a proportionate share of Mill's counterparts today (college professors, literary critics, journalists, and pundits). Plainly, there are many millions of people who have indeed experienced both higher and lower pleasures, and continue to enjoy both.

So Mill did not have a very good psychological theory of "entertainment", even if he had a good aspirational philosophy of living. I doubt that Netflix would want JS Mill to serve as programming chief for its upcoming seasons.

There is probably a nugget of truth in the idea that the challenge and complexity of an activity is a dimension of its continuing appeal to the individuals who engage in the activity. Perhaps Go masters and chess masters take greater satisfaction in their games than do checkers players and aficionados of tic-tac-toe, and people who love literature may take more satisfaction from a novel by Boris Pasternak than Nora Roberts because of the relative complexity, unpredictability, and nuance of Pasternak's love story. But it is clear that pleasure in entertainment derives from more than this: humor, suspense, nostalgia, complicated plots, character development, amazing special effects, depictions of human emotions, violence, plots that speak to one's own experience, and so on indefinitely. Why these elements produce pleasure, however, is still unclear.

Perhaps we need to go back to ancient Greek philosophers in our quest for the sources of pleasure in "entertainment". Is Aristotle right that we enjoy tragedy because of the catharsis it provides, and comedy, because of the physical pleasure we take in laughter? Does Epicurus have anything to tell us about why we enjoy watching *Breaking Bad* and *Blazing Saddles*? Can the Stoics shed light on why *Santa Claus Conquers the Martians* (1964) and *Heaven's Gate* (1980) were so very unpopular with the movie audience? Does Plato have anything to say about whether *Waiting for Godot* is really entertaining, and whether people can take pleasure in viewing it?

Chapter 5

TECHNOLOGY FAILURE AND RISK

Section 1

TECHNOLOGY FAILURE

My current research interests have to do with organizational dysfunction and large-scale technology failures. I am interested in probing the ways in which organizational failures and dysfunctions have contributed to large accidents like Bhopal, Fukushima, and the Deepwater Horizon disaster. I've had to confront an important question in taking on this research interest: what can philosophy bring to the topic that would not be better handled by engineers, organizational specialists, or public policy experts?

One answer is the diversity of viewpoint that a philosopher can bring to the discussion. It is evident that technology failures invite analysis from all of these specialized experts, and more. But there is room for productive contribution from reflective observers who are not committed to any of these disciplines. Philosophers have a long history of taking on big topics outside the defined canon of "philosophical problems", and often those engagements have proven fruitful. In this particular instance, philosophy can look at organizations and technology in a way that is more likely to be interdisciplinary, and perhaps can help to see dimensions of the problem that are less apparent from a purely disciplinary perspective.

There is also a rationale based on the terrain of the philosophy of science. Philosophers of biology have usually attempted to learn as much about the science of biology as they can manage, but they lack the level of expertise of a research biologist, and it is rare for a philosopher to make an original contribution to the scientific biological literature. Nonetheless it is clear that philosophers have a great deal to add to scientific research in biology. They can contribute to better reasoning about the implications of various theories, they can probe the assumptions about confirmation and explanation that are in use, and they can contribute to important conceptual disagreements. Biology is in a better state because of the work of philosophers like David Hull and Elliot Sober.

Philosophers have also made valuable contributions to science and technology studies, bringing a viewpoint that incorporates insights from the philosophy of science and a sensitivity to the social groundedness of technology. STS studies have proven to be a fruitful place for interaction between historians, sociologists, and philosophers. Here again, the concrete study of the causes and context of large technology failure may be assisted by a philosophical perspective.

There is also a normative dimension to these questions about technology failure for which philosophy is well prepared. Accidents hurt people, and sometimes the causes of accidents involve culpable behavior by individuals and corporations. Philosophers have a long history of contribution to these kinds of problems of fault, law, and just management of risks and harms.

Finally, it is realistic to say that philosophy has an ability to contribute to social theory. Philosophers can offer imagination and critical attention to the problem of creating new conceptual schemes for understanding the social world. This capacity seems relevant to the problem of describing, analyzing, and explaining largescale failures and disasters.

The situation of organizational studies and accidents is in some ways more hospitable for contributions by a philosopher than other "wicked problems" in the world around us. An accident is complicated and complex but not particularly obscure. The field is unlike quantum mechanics or climate dynamics, which are inherently difficult for non-specialists to understand. The challenge with accidents is to identify a multi-layered analysis of the causes of the accident that permits observers to have a balanced and operative understanding of the event. And this is a situation where the philosopher's perspective is most useful. We can offer higher-level descriptions of the relative importance of different kinds of causal factors. Perhaps the role here is analogous to messenger RNA, providing a cross-disciplinary kind of communications flow. Or it is analogous to the role of philosophers of history who have offered gentle critique of the cliometrics school for its over-dependence on a purely statistical approach to economic history.

So it seems reasonable enough for a philosopher to attempt to contribute to this set of topics, even if the disciplinary expertise a philosopher brings is more weighted towards conceptual and theoretical discussions than undertaking original empirical research in the domain.

What I expect to be the central finding of this research is the idea that a pervasive and often unrecognized cause of accidents is a systemic organizational defect of some sort, and that it is enormously important to have a better understanding of common forms of these deficiencies. This is a bit analogous to a paradigm shift in the study of accidents. And this view has important policy implications. We can make disasters less frequent by improving the organizations through which technology processes are designed and managed.

Section 2

EXPLAINING TECHNOLOGICAL FAILURE

Technology failure is often spectacular and devastating -- witness Bhopal, Three Mile Island, Chernobyl, the Challenger disaster, and the DC10 failures of the 1970s. But in addition to being a particularly important cause of human suffering, technology failures are often very complicated social outcomes that involve a number of different kinds of factors. And this makes them interesting topics for social science study.

It is fairly common to attribute spectacular failures to a small number of causes -- for example, faulty design, operator error, or a conjunction of unfortunate but singly non-fatal accidents. What sociologists who have studied technology failures have been able to add is the fact that the root causes of disastrous failures can often be traced back to deficiencies of the social organizations in which they are designed, used, or controlled (Charles Perrow, *Normal Accidents: Living with High-Risk Technologies*). Technology failures are commonly the result of specific social organizational defects; so technology failure is often or usually a social outcome, not simply a technical or mechanical misadventure. (Dietrich Dorner's *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations* is a fascinating treatment of a number of cases of failure; Eliot Cohen's *Military Misfortunes: The Anatomy of Failure in War* provides an equally interesting treatment of military failures; for example, the American failure to suppress submarine attacks on merchant shipping off the US coast in the early part of World War II.)

First, a few examples. The Challenger space shuttle was destroyed as a result of O-rings in the rocket booster units that became brittle because of the low launch temperature -- evidently an example of faulty design. But various observers have asked the more fundamental question: what features of the science-engineering-launch command process that was in place within NASA and between NASA and its aerospace suppliers led it to break down so profoundly (Diane Vaughan, *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*)? What organizational defects made it possible for this extended group of talented scientists and engineers to come to the decision to launch over the specific warnings that were brought forward by the rocket provider's team about the danger of a cold-temperature launch? Edward Tufte attributes the failure to poor scientific communication (*Visual Explanations: Images and Quantities, Evidence and Narrative*); Morton Thiokol engineer Roger Boisjoly attributes it to an excessively hierarchical and deferential relation between the engineers and the launch decision-makers. Either way, features of the NASA decision-making process -- social-organizational features -- played a critical role.

Bhopal represents another important case. Catastrophic failure of a Union Carbide pesticide plant in Bhopal, India in 1984 led to a release of a highly toxic gas. The toxic cloud passed into the densely populated city of Bhopal. Half a million people were affected, and between 16 and 30 thousand people died as a result. A chemical plant is a complex physical system. But even more, it is operated and maintained by a complex social organization, involving training, supervision, and operational assessment and oversight. In his careful case study of Bhopal, Paul Shrivastava maintains that this disaster was caused by a set of persistent and recurring organizational failures, especially in the areas of training and supervision of operators (*Bhopal: Anatomy of Crisis*).

Close studies of the nuclear disasters at Chernobyl and Three Mile Island have been equally fruitful in terms of shedding light on the characteristics of social, political, and business organization that have played a role in causing these great disasters. The stories are different in the two cases; but in each case, it turns out that social factors, including both organizational features internal to the nuclear plants and political features in the surrounding environment, played a role in the occurrence and eventual degree of destruction associated with the disasters.

These cases illustrate several important points. First, technology failures and disasters almost always involve a crucial social dimension -- in the form of the organizations and systems through which the technology is developed, deployed, and maintained and the larger social environment within which the technology is situated. Technology systems are social systems. Second, technology failures therefore constitute an important subject matter for sociological and organizational research. Sociologists can shed light on the ways in which a complex technology might fail. And third, and most importantly, the design of safe systems -- particularly systems that have the potential for creating great harms -- needs to be an interdisciplinary effort. The perspectives of sociologists and organizational theorists need to be incorporated as deeply as those of industrial and systems engineers into the design of systems that will preserve a high degree of safety. This is an important realization for the high profile risky industries -- aviation, chemicals, nuclear power. But it is also fundamental for other important social institutions, including especially hospitals and health systems. Safe technologies will only exist when they are embedded in safe, fault-tolerant organizations and institutions. And all of this means, in turn, that there is an urgent need for a sociology of safety.

Section 3

TECHNOLOGY ACCIDENTS

Complex socio-technical systems fail; that is, accidents occur. And it is enormously important for engineers and policy makers to have a better way of thinking about accidents than is the current protocol following an air crash, a chemical plant fire, or the release of a contaminated drug. We need to understand better what the systems and organizational causes of an accident are; even more importantly, we need to have a basis for improving the safe functioning of complex socio-technical systems by identifying better processes and better warning indicators of impending failure.

A long-term leader in the field of systems-safety thinking is Nancy Leveson, a professor of aeronautics and astronautics at MIT and the author of *Safeware: System Safety and Computers* (1995) and *Engineering a Safer World: Systems Thinking Applied to Safety* (2012). Leveson has been a particular advocate for two insights: looking at safety as a systems characteristic, and looking for the organizational and social components of safety and accidents as well as the technical event histories that are more often the focus of accident analysis. Her approach to safety and accidents involves looking at a technology system in terms of the set of controls and constraints that have been designed into the process to prevent accidents. "Accidents are seen as resulting from inadequate control or enforcement of constraints on safety-related behavior at each level of the system development and system operations control structures." (25)

The abstract for her essay "A New Accident Model for Engineering Safety" captures both points.

New technology is making fundamental changes in the etiology of accidents and is creating a need for changes in the explanatory mechanisms used. We need better and less subjective understanding of why accidents occur and how to prevent future ones. The most effective models will go beyond assigning blame and instead help engineers to learn as much as possible about all the factors involved, including those related to social and organizational structures. This paper presents a new accident model founded on basic systems theory con-

cepts. The use of such a model provides a theoretical foundation for the introduction of unique new types of accident analysis, hazard analysis, accident prevention strategies including new approaches to designing for safety, risk assessment techniques, and approaches to designing performance monitoring and safety metrics.

The accident model she describes in this article and elsewhere is STAMP (Systems-Theoretic Accident Model and Processes). Here is a short description of the approach.

In STAMP, systems are viewed as interrelated components that are kept in a state of dynamic equilibrium by feedback loops of information and control. A system in this conceptualization is not a static design—it is a dynamic process that is continually adapting to achieve its ends and to react to changes in itself and its environment. The original design must not only enforce appropriate constraints on behavior to ensure safe operation, but the system must continue to operate safely as changes occur. The process leading up to an accident (loss event) can be described in terms of an adaptive feedback function that fails to maintain safety as performance changes over time to meet a complex set of goals and values.... The basic concepts in STAMP are constraints, control loops and process models, and levels of control. (12)

The other point of emphasis in Leveson's treatment of safety is her consistent effort to include the social and organizational forms of control that are a part of the safe functioning of a complex technological system.

Event-based models are poor at representing systemic accident factors such as structural deficiencies in the organization, management deficiencies, and flaws in the safety culture of the company or industry. An accident model should encourage a broad view of accident mechanisms that expands the investigation from beyond the proximate events. (6)

She treats the organizational backdrop of the technology process in question as being a crucial component of the safe functioning of the process.

Social and organizational factors, such as structural deficiencies in the organization, flaws in the safety culture, and inadequate management decision making and control are directly *represented in the model and treated as complex processes rather than simply modeling their reflection in an event chain. (26)*

And she treats organizational features as another form of control system (along the lines of Jay Forrester's early definitions of systems in Industrial Dynamics).

Modeling complex organizations or industries using system theory involves dividing them into hierarchical levels with control processes operating at the interfaces between levels (Rasmussen, 1997). Figure 4 shows a generic socio-technical control model. Each system, of course, must be modeled to reflect its specific features, but all will have a structure that is a variant on this one. (17)

Here is figure 4:

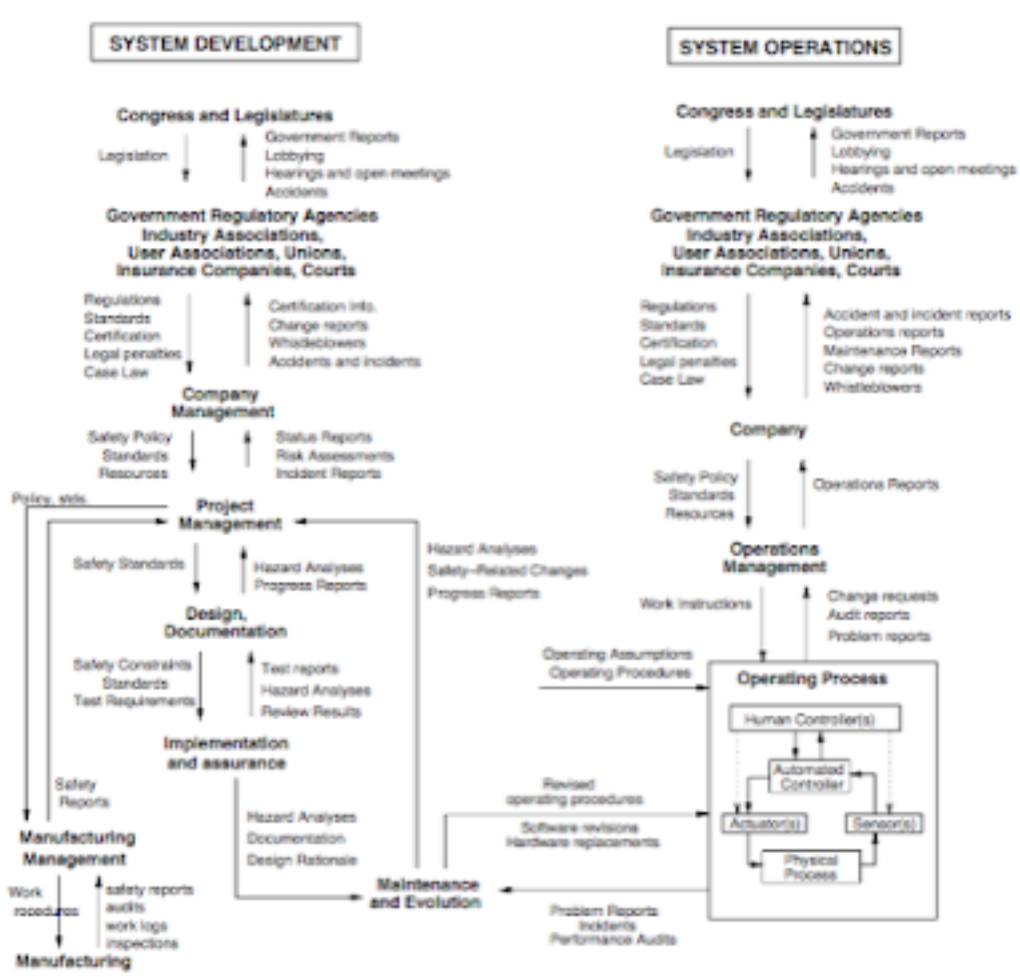


Figure 4: General Form of a Model of Socio-Technical Control.

The approach embodied in the STAMP framework is that safety is a systems effect, dynamically influenced by the control systems embodied in the total process in question.

In STAMP, systems are viewed as interrelated components that are kept in a state of dynamic equilibrium by feedback loops of information and control. A system in this conceptualization is not a static design—it is a dynamic process that is continually adapting to achieve its ends and to react to changes in itself and its environment. The original design must not only enforce appropriate constraints on behavior to ensure safe operation, but the system must continue to operate safely as changes occur. The process leading up to an accident (loss event) can be described in terms of an adaptive feedback function that fails to maintain safety as performance changes over time to meet a complex set of goals and values. (12)

And:

In systems theory, systems are viewed as hierarchical structures where each level imposes constraints on the activity of the level beneath it—that is, constraints or lack of constraints at a higher level allow or control lower-level behavior (Checkland, 1981). Control laws are constraints on the relationships between the values of system variables. Safety-related control laws or constraints therefore specify those relationships between system variables that constitute the nonhazardous system states, for example, the power must never be on when the access door is open. The control processes (including the physical design) that enforce these constraints will limit system behavior to safe changes and adaptations. (17)

Leveson's understanding of systems theory brings along with it a strong conception of "emergence". She argues that higher levels of systems possess properties that cannot be reduced to the properties of the components, and that safety is one such property:

In systems theory, complex systems are modeled as a hierarchy of levels of organization, each more complex than the one below, where a level is characterized by having emergent or irreducible properties. Hierarchy theory deals with the fundamental differences between one level of complexity and another. Its ultimate aim is to explain the relationships between different levels: what generates the levels, what separates them, and what links them. Emergent *properties associated with a set of components at one level in a hierarchy are related to constraints upon the degree of freedom of those components. (11)*

But her understanding of "irreducible" seems to be different from that commonly used in the philosophy of science. She does in fact believe that these higher-level properties can be explained by the system of properties at the lower levels -- for example, in this passage she asks "... what generates the levels" and how the emergent properties are "related to constraints" imposed on the lower levels. In other words, her position seems to be similar to that advanced by Dave Elder-Vass: emergent properties are properties at a higher level that are not possessed by the components, but which depend upon the interactions and composition of the lower-level components.

The domain of safety engineering and accident analysis seems like a particularly suitable place for Bayesian analysis. It seems unavoidable that accident analysis involves both frequency-based probabilities (e.g. the frequency of pump failure) and expert-based estimates of the likelihood of a particular kind of failure (e.g. the likelihood that a train operator will slacken attention to track warnings in response to company pressure on timetable). Bayesian techniques are suitable for the task of combining these various kinds of estimates of risk into a unified calculation.

The topic of safety and accidents is particularly relevant to the philosophy of technology because it expresses very clearly the causal complexity of the social world in which we live. And rather than simply ignoring that complexity, the systematic study of accidents gives us an avenue for arriving at better ways of representing, modeling, and intervening in parts of that complex world.

Section 4

SYSTEM SAFETY

An ongoing thread of posts here is concerned with organizational causes of large technology failures. The driving idea is that failures, accidents, and disasters usually have a dimension of organizational causation behind them. The corporation, research office, shop floor, supervisory system, intra-organizational information flow, and other social elements often play a key role in the occurrence of a gas plant fire, a nuclear power plant malfunction, or a military disaster. There is a tendency to look first and foremost for one or more individuals who made a mistake in order to explain the occurrence of an accident or technology failure; but researchers such as Perrow, Vaughan, Tierney, and Hopkins have demonstrated in detail the importance of broadening the lens to seek out the social and organizational background of an accident.

It seems important to distinguish between system flaws and organizational dysfunction in considering all of the kinds of accidents mentioned here. We might specify system safety along these lines. Any complex process has the potential for malfunction. Good system design means creating a flow of events and processes that make accidents inherently less likely. Part of the task of the designer and engineer is to identify chief sources of harm inherent in the process -- release of energy, contamination of food or drugs, unplanned fission in a nuclear plant -- and design fail-safe processes so that these events are as unlikely as possible. Further, given the complexity of contemporary technology systems it is critical to attempt to anticipate unintended interactions among subsystems -- each of which is functioning correctly but that lead to disaster in unusual but possible interaction scenarios.

In a nuclear processing plant, for example, there is the hazard of radioactive materials being brought into proximity with each other in a way that creates unintended critical mass. Jim Mahaffey's *Atomic Accidents: A History of Nuclear Meltdowns and Disasters: From the Ozark Mountains to Fukushima* offers numerous examples of such unintended events, from the careless handling of plutonium scrap in a machining process to the transfer of a fissionable liquid from a vessel of one shape to another. We might try to handle these risks as an organizational problem: more and better training for operatives about the importance of handling nuclear materials according to established protocols, and effective supervision and oversight to ensure that the protocols are observed on a regular basis. But it is also possible to design the material processes within a nuclear plant in a way that makes unintended criticality virtually impossible -- for example, by storing radioactive solutions in containers that simply cannot be brought into close proximity with each other.

Nancy Leveson is a national expert on defining and applying principles of system safety. Her book *Engineering a Safer World: Systems Thinking Applied to Safety* is a thorough treatment of her thinking about this subject. She offers a handful of compelling reasons for believing that safety is a system-level characteristic that requires a systems approach: the fast pace of technological change, reduced ability to learn from experience, the changing nature of accidents, new types of hazards, increasing complexity and coupling, decreasing tolerance for single accidents, difficulty in selecting priorities and making tradeoffs , more complex relationships between humans and automation, and changing regulatory and public view of safety (kl 130 ff.). Particularly important in this list is the comment about complexity and coupling: "The operation of some systems is so complex that it defies the understanding of all but a few experts, and sometimes even they have incomplete information about the system's potential behavior" (kl 137).

Given the fact that safety and accidents are products of whole systems, she is critical of the accident methodology generally applied to serious industrial, aerospace, and chemical accidents. This methodology involves tracing the series of events that led to the outcome, and identifying one or more events as the critical cause of the accident. However, she writes:

In general, event-based models are poor at representing systemic accident factors such as structural deficiencies in the organization, management decision making, and flaws in the safety culture of the or industry. An accident model should encourage a broad view of accident mechanisms that expands the investigation beyond the proximate events. A narrow focus on technological components and pure engineering activities or a similar narrow focus *on operator errors may lead to ignoring some of the most important factors in terms of preventing future accidents. (kl 452)*

Here is a definition of system safety offered later in ESW in her discussion of the emergence of the concept within the defense and aerospace fields in the 1960s:

System Safety ... is a subdiscipline of system engineering. It was created at the same time and for the same reasons. The defense community tried using the standard safety engineering techniques on their complex new systems, but the limitations became clear when interface and component interaction problems went unnoticed until it was too late, resulting in many losses and near misses. When these early aerospace accidents were investigated, the causes of a large percentage of them were traced to deficiencies in design, operations, and management. Clearly, big changes were needed. System engineering along with its subdiscipline, System Safety, were developed to tackle these problems. (kl 1007)

Here Leveson mixes system design and organizational dysfunctions as system-level causes of accidents. But much of her work in this book and her earlier Safeware: System Safety and Computers gives extensive attention to the design faults and component interactions that lead to accidents -- what we might call system safety in the narrow or technical sense.

A systems engineering approach to safety starts with the basic assumption that some properties of systems, in this case safety, can only be treated adequately in the context of the social and technical system as a whole. A basic assumption of systems engineering is that optimization of individual components or subsystems will not in general lead to a system optimum; in fact, improvement of a particular subsystem may actually worsen the overall system performance because of complex, nonlinear interactions among the components. (kl 1007)

Overall, then, it seems clear that Leveson believes that both organizational features and technical system characteristics are part of the systems that created the possibility for accidents like Bhopal, Fukushima, and Three Mile Island. Her own accident model designed to help identify causes of accidents, STAMP (Systems-Theoretic Accident Model and Processes) emphasizes both kinds of system properties.

Using this new causality model ... changes the emphasis in system safety from preventing failures to enforcing behavioral safety constraints. Component failure accidents are still included, but or conception of causality is extended to include component interaction accidents. Safety is reformulated as a control problem rather than a reliability problem. (kl 1062)

In this framework, understanding why an accident occurred requires determining why the control was ineffective. Preventing future accidents requires shifting from a focus on preventing failures to the broader goal of designing and implementing controls that will enforce the necessary constraints. (kl 1084)

Leveson's brief analysis of the Bhopal disaster in 1984 (kl 384 ff.) emphasizes the organizational dysfunctions that led to the accident -- and that were completely ignored by the Indian state's accident investigation of the accident: out-of-service gauges, alarm deficiencies, inadequate response to prior safety audits, shortage of oxygen masks, failure to inform the police or surrounding community of the accident, and an environment of cost cutting that impaired maintenance and staffing. "When all the factors, including indirect and systemic ones, are considered, it becomes clear that the maintenance worker was, in fact, only a minor and somewhat irrelevant player in the loss. Instead, degradation in the safety margin occurred over time and without any particular single decision to do so but simply as a series of decisions that moved the plant slowly toward a situation where any slight error would lead to a major accident" (kl 447).

Section 5

HIGH-RELIABILITY ORGANIZATIONS

December 28, 2019

https://understandingsociety.blogspot.com/2019/12/high-reliability-organizations.html

Charles Perrow takes a particularly negative view of the possibility of safe management of high-risk technologies in *Normal Accidents: Living with High-Risk Technologies*. His summary of the Three Mile Island accident is illustrative: “The system caused the accident, not the operators” (12). Perrow’s account of TMI is chiefly an account of complex and tightly-coupled system processes, and the difficulty these processes create for operators and managers when they go wrong. And he is doubtful that the industry can safely manage its nuclear plants.

It is interesting to note that systems engineer and safety expert Nancy Leveson addresses the same features of “system accidents” that Perrow addresses, but with a greater level of confidence about the possibility of creating engineering and organizational enhancements. A recent expression of her theory of technology safety is provided in *Engineering a Safer World: Systems Thinking Applied to Safety* and *Resilience Engineering: Concepts and Precepts*.

In examining the safety of high-risk industries, our goal should be to identify some of the behavioral, organizational, and regulatory dysfunctions that increase the likelihood and severity of accidents, and to consider organizational and behavioral changes that would serve to reduce the risk and severity of accidents. This is the approach taken by a group of organizational theorists, engineers, and safety experts who explore the idea and practice of a "high reliability organization". Scott Sagan describes the HRO approach in these terms in *The Limits of Safety*:

> *The common assumption of the high reliability theorists is not a naive belief in the ability of human beings to behave with perfect rationality, it is the much more plausible belief that organizations, properly designed and managed, can compensate for well-known human frailties and can therefore be significantly more rational and effective than can individuals. (Sagan, 16)*

Sagan lists several conclusions advanced by HRO theorists, based on a small number of studies of high-risk organizational environments. Researchers have identified a set of organizational features that appear to be common among HROs:

- Leadership safety objectives: priority on avoiding altogether serious operational failures
- Organizational leaders must place high priority on safety in order to communicate this objective clearly and consistently to the rest of the organization
- The need for redundancy. Multiple and independent channels of communication, decision-making, and implementation can produce a highly reliable overall system
- Decentralization -- authority must exist in order to permit rapid and appropriate responses to dangers by individuals closest to the problems

- culture – recruit individuals who help maintain a strong organizational culture emphasizing safety and reliability
- continuity – maintain continuous operations, vigilance, and training
- organizational learning – learn from prior accidents and near-misses.
- Improve the use of simulation and imagination of failure scenarios

Here is Sagan's effort to compare Normal Accident Theory with High Reliability Organization Theory:

46 CHAPTER 1

TABLE 1.1
Competing Perspectives on Safety with Hazardous Technologies

High Reliability Theory	*Normal Accidents Theory*
Accidents can be prevented through good organizational design and management.	Accidents are inevitable in complex and tightly coupled systems.
Safety is the priority organizational objective.	Safety is one of a number competing objectives.
Redundancy enhances safety: duplication and overlap can make "a reliable system out of unreliable parts."	Redundancy often causes accidents: it increases interactive complexity and opaqueness and encourages risk-taking.
Decentralized decision-making is needed to permit prompt and flexible field-level responses to surprises.	Organizational contradiction: decentralization is needed for complexity, but centralization is needed for tightly coupled systems.
A "culture of reliability" will enhance safety by encouraging uniform and appropriate responses by field-level operators.	A military model of intense discipline, socialization, and isolation is incompatible with democratic values.
Continuous operations, training, and simulations can create and maintain high reliability operations.	Organizations cannot train for unimagined, highly dangerous, or politically unpalatable operations.
Trial and error learning from accidents can be effective, and can be supplemented by anticipation and simulations.	Denial of responsibility, faulty reporting, and reconstruction of history cripples learning efforts.

The genuinely important question here is whether there are indeed organizational arrangements, design principles, and behavioral practices that are consistently effective in significantly reducing the incidence and harmfulness of accidents in high-risk enterprises, or whether on the other hand, the ideal of a "High Reliability Organization" is more chimera than reality.

A respected organizational theorist who has written on high-reliability organizations and practices extensively is Karl Weick. He and Kathleen Sutcliffe attempt to draw some useable maxims for high reliability in *Managing the Unexpected: Sustained Performance in a Complex World*. They use several examples of real-world business failures to illustrate their central recommendations, including an in-depth case study of the Washington Mutual financial collapse in 2008.

The chief recommendations of their book come down to five maxims for enhancing reliability:

1. Pay attention to weak signals of unexpected events
2. Avoid extreme simplification
3. Pay close attention to operations
4. Maintain a commitment to resilience
5. Defer to expertise

Maxim 1 (preoccupation with failure) encourages a style of thinking -- an alertness to unusual activity or anomalous events and a commitment to learning from near-misses in the past. This alertness is both individual and organizational; individual members of the organization need to be alert to weak signals in their areas, and managers need to be receptive to hearing the "bad news" when ominous signals are reported. By paying attention to "weak signals" of possible failure, managers will have more time to design solutions to failures when they emerge.

Maxim 2 addresses the common cognitive mistake of subsuming unusual or unexpected outcomes under more common and harmless categories. Managers should be reluctant to accept simplifications. The Columbia space shuttle disaster seems to fall in this category, where senior NASA managers dismissed evidence of foam strike during lift-off by subsuming it under many earlier instances of debris strikes.

Maxim 3 addresses the organizational failure associated with distant management -- top executives who are highly "hands-off" in their knowledge and actions with regard to ongoing operations of the business. (The current Boeing story seems to illustrate this failure; even the decision to move the corporate headquarters to Chicago, very distant from the engineering and manufacturing facilities in Seattle, illustrates a hands-off attitude towards operations.) Executives who look at their work as "the big picture" rather than ensuring high-quality activity within the actual operations of the organization are likely to oversee disaster at some point.

Maxim 4 is both cognitive and organizational. "Resilience" refers to the "ability of an organization (system) to maintain or regain a dynamically stable state, which allows it to continue operations after a major mishap and/ or in the presence of a continuous stress". A resilient organization is one where process design has been carried out in order to avoid single-point failures, where resources and tools are available to address possible "off-design" failures, and where the interruption of one series of activities (electrical power) does not completely block another vital series of activities (flow of cooling water). A resilient team is one in which multiple capable individuals are ready to work together to solve problems, sometimes in novel ways, to ameliorate the consequences of unexpected failure.

Maxim 5 emphasizes the point that complex activities and processes need to be managed by teams incorporating experience, knowledge, and creativity in order to be able to confront and surmount unexpected failures. Weick and Sutcliffe give telling examples of instances where key expertise was lost at the frontline level through attrition or employee discouragement, and where senior executives substituted their judgment for the recommendations of more expert subordinates.

These maxims involve a substantial dose of cognitive practice, changing the way that employees, managers, and executives think: the importance of paying attention to signs of unexpected outcomes (pumps that repeatedly fail in a refinery), learning from near-misses, making full use of the expertise of members of the organization, ... It is also possible to see how various organizations could be evaluated in terms of their performance on these five maxims -- before a serious failure has occurred -- and could improve performance accordingly.

It is interesting to observe, however, that Weick and Sutcliffe do not highlight some factors that have been given strong priority in other treatments of high-reliability organizations: the importance of establishing a high priority for system safety in the highest management levels of the organization (which unavoidably competes with cost and profit pressures), the organizational feature of an empowered safety executive outside the scope of production and business executives in the organization, the possible benefits of a somewhat decentralized system of control, the possible benefits of redundancy, the importance of well-designed training aimed at enhancing system safety as well as personal safety, and the importance of creating a culture of honesty and compliance when it comes to safety. When mid-level managers are discouraged from bringing forward their concerns about the "signals" they perceive in their areas, this is a pre-catastrophe situation.

There is a place in the management literature for a handbook of research on high-reliability organizations; at present, such a resource does not exist.

(See also Sagan and Blanford's volume *Learning from a Disaster: Improving Nuclear Safety and Security after Fukushima*.)

Section 6

THE DEEPWATER HORIZON DISASTER

The Deepwater Horizon oil rig explosion, fire, and uncontrolled release of oil into the Gulf of Mexico is a disaster of unprecedented magnitude. This disaster in the Gulf of Mexico appears to be more serious in objective terms than the Challenger space shuttle disaster in 1986 -- in terms both of immediate loss of life and in terms of overall harm created. And sadly, it appears likely that the accident will reveal equally severe failures of management of enormously hazardous processes, defects in the associated safety engineering analysis, and inadequacies of the regulatory environment within which the activity took place. The Challenger disaster fundamentally changed the ways that we thought about safety in the aerospace field. It is likely that this disaster too will force radical new thinking and new procedures concerning how to deal with the inherently dangerous processes associated with deep-ocean drilling.

Nancy Leveson is an important expert in the area of systems safety engineering, and her book, *Safeware: System Safety and Computers*, is a genuinely important contribution. Leveson led the investigation of the role that software design might have played in the Challenger disaster (link). Here is a short, readable white paper of hers on system safety engineering (link) that is highly relevant to the discussions that will need to occur about deep-ocean drilling. The paper does a great job of laying out how safety has been analyzed in several high-hazard industries, and presents a set of basic principles for systems safety design. She discusses aviation, the nuclear industry, military aerospace, and the chemical industry; and she points out some important differences across industries when it comes to safety engineering. Here is an instructive description of the safety situation in military aerospace in the 1950s and 1960s:

Within 18 months after the fleet of 71 Atlas F missiles became operational, four blew up in their silos during operational testing. The missiles also had an extremely low launch success rate. An Air Force manual describes several of these accidents:

An ICBM silo was destroyed because the counterweights, used to balance the silo elevator on the way up and down in the silo, were designed with consideration only to raising a fueled missile to the surface for firing. There was no consideration that, when you were not firing in anger, you had to bring the fueled missile back down to defuel.

The first operation with a fueled missile was nearly successful. The drive mechanism held it for all but the last five feet when gravity took over and the missile dropped back. Very suddenly, the 40-foot diameter silo was altered to about 100-foot diameter.

During operational tests on another silo, the decision was made to continue a test against the safety engineer's advice when all indications were that, because of high oxygen concentrations in the silo, a catastrophe was imminent. The resulting fire destroyed a missile and caused extensive silo damage. In another accident, five people were killed when a single-point failure in a hydraulic system caused a 120-ton door to fall.

Launch failures were caused by reversed gyros, reversed electrical plugs, bypass of procedural steps, and by management decisions to continue, in spite of contrary indications, because of schedule pressures. (from the Air Force System Safety Handbook for Acquisition Managers, Air Force Space Division, January 1984)

Leveson's illustrations from the history of these industries are fascinating. But even more valuable are the principles of safety engineering that she recapitulates. These principles seem to have many implications for deep-ocean drilling and associated technologies and systems. Here is her definition of systems safety:

System safety uses systems theory and systems engineering approaches to prevent foreseeable accidents and to minimize the result of unforeseen ones. Losses in general, not just human death or injury, are considered. Such losses may include destruction of property, loss of mission, and environmental harm. The primary concern of system safety is the management of hazards: their identification, evaluation, elimination, and control through analysis, design and management procedures.

Here are several fundamental principles of designing safe systems that she discusses:

- System safety emphasizes building in safety, not adding it on to a completed design.
- System safety deals with systems as a whole rather than with subsystems or components.
- System safety takes a larger view of hazards than just failures.
- System safety emphasizes analysis rather than past experience and standards.
- System safety emphasizes qualitative rather than quantitative approaches.
- Recognition of tradeoffs and conflicts.
- System safety is more than just system engineering.

And here is an important summary observation about the complexity of safe systems:

> *Safety is an emergent property that arises at the system level when components are operating together. The events leading to an accident may be a complex combination of equipment failure, faulty maintenance, instrumentation and control problems, human actions, and design errors. Reliability analysis considers only the possibility of accidents related to failures; it does not investigate potential damage that could result from successful operation of the individual components.*

How do these principles apply to the engineering problem of deep-ocean drilling? Perhaps the most important implications are these: a safe system needs to be based on careful and comprehensive analysis of the hazards that are inherently involved in the process; it needs to be designed with an eye to handling those hazards safely; and it can't be done in a piecemeal, "fly-test-fly" fashion.

It would appear that deep-ocean drilling is characterized by too little analysis and too much confidence in the ability of engineers to "correct" inadvertent outcomes ("fly-fix-fly"). The accident that occurred in the Gulf last month can be analyzed into several parts. First is the explosion and fire that destroyed the drilling rig and led to the tragic loss of life of 11 rig workers. And the second is the uncalculated harms caused by the uncontrolled venting of perhaps a hundred thousand barrels of crude oil to date into the Gulf of Mexico, now threatening the coasts and ecologies of several states. Shockingly, there is now no high-reliability method for capping the well at a depth of over 5,000 feet; so the harm can continue to worsen for a very extended period of time.

The safety systems on the platform itself will need to be examined in detail. But the bottom line will probably look something like this: the platform is a complex system vulnerable to explosion and fire, and there was always a calculable (though presumably small) probability of catastrophic fire and loss of the ship. This is pretty analogous to the problem of safety in aircraft and other complex electro-mechanical systems. The loss of life in the incident is terrible but confined. Planes crash and ships sink.

What elevates this accident to a globally important catastrophe is what happened next: destruction of the pipeline leading from the wellhead 5,000 feet below sea level to containers on the surface; and the failure of the shutoff valve system on the ocean floor. These two failures have resulted in unconstrained release of a massive and uncontrollable flow of crude oil into the Gulf and the likelihood of environmental harms that are likely to be greater than the Exxon Valdez.

Oil wells fail on the surface, and they are difficult to control. But there is a well-developed technology that teams of oil fire specialists like Red Adair employ to cap the flow and end the damage. We don't have anything like this for wells drilled under water at the depth of this incident; this accident is less accessible than objects in space for corrective intervention. So surface well failures conform to a sort of epsilon-delta relationship: an epsilon accident leads to a limited delta harm. This deep-ocean well failure in the Gulf is catastrophically different: the relatively small incident on the surface is resulting in an unbounded and spiraling harm.

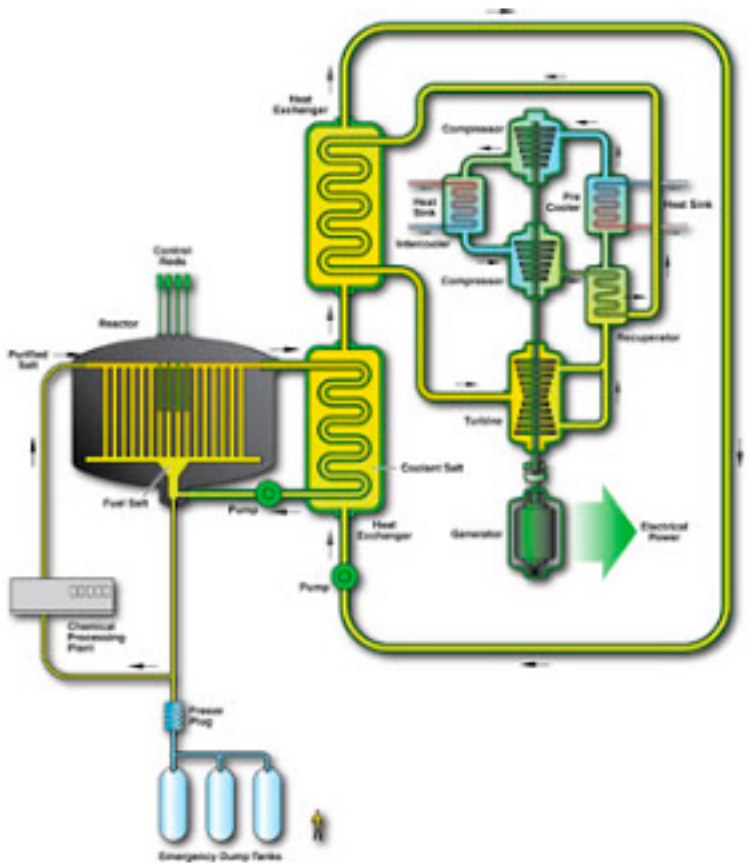
So was this a foreseeable hazard? Of course it was. There was always a finite probability of total loss of the platform, leading to destruction of the pipeline. There was also a finite probability of failure of the massive sea-floor emergency shutoff valve. And, critically, it was certainly known that there is no high-reliability fix in the event of failure of the shutoff valve. The effort to use the dome currently being tried by BP is untested and unproven at this great depth. The alternative of drilling a second well to relieve pressure may work; but it will take weeks or months. So essentially, when we reach the end of this failure pathway, we arrive at this conclusion: catastrophic, unbounded failure. If you reach this point in the fault tree, there is almost nothing to be done. And this is a totally irrational outcome to tolerate; how could any engineer or regulatory agency have accepted the circumstances of this activity, given that one possible failure pathway would lead predictably to unbounded harms?

There is one line of thought that might have led to the conclusion that deep ocean drilling is acceptably safe: engineers and policy makers might have optimistically overestimated the reliability of the critical components. If we estimate that the probability of failure of the platform is 1/1000, failure of the pipeline is 1/100, and failure of the emergency shutoff valve is 1/10,000 -- then one might say that the probability of the nightmare scenario is vanishingly small: one in a billion. Perhaps one might reason that we can disregard scenarios with this level of likelihood. Reasoning very much like this was involved in the original safety designs of the shuttle (*Safeware: System Safety and Computers*). But several things are now clear: this disaster was not virtually impossible. In fact, it actually occurred. And second, it seems likely enough that the estimates of component failure are badly understated.

What does this imply about deep ocean drilling? It seems inescapable that the current state of technology does not permit us to take the risk of this kind of total systems failure. Until there is a reliable and reasonably quick technology for capping a deep-ocean well, the small probability of this kind of failure makes the use of the technology entirely unjustifiable. It makes no sense at all to play Russian roulette when the cost of failure is massive and unconstrained ecological damage.

Section 7

TECHNOLOGY LOCK-IN ACCIDENTS



Organizational and regulatory features are sometimes part of the causal background of important technology failures. This is particularly true in the history of nuclear power generation. The promise of peaceful uses of atomic energy was enormously attractive at the end of World War II. In abstract terms the possibility of generating useable power from atomic reactions was quite simple. What was needed was a controllable fission reaction in which the heat produced by fission could be captured to run a steam-powered electrical generator.

The technical challenges presented by harnessing nuclear fission in a power plant were large. Fissionable material needed to be produced as useable fuel sources. A control system needed to be designed to maintain the level of fission at a desired level. And, most critically, a system for removing heat from the fissioning fuel needed to be designed so that the reactor core would not overheat and melt down, releasing energy and radioactive materials into the environment.

Early reactor designs took different approaches to the heat-removal problem. Liquid metal reactors used a metal like sodium as the fluid that would run through the core removing heat to a heat sink for dispersal; and water reactors used pressurized water to serve that function. The sodium breeder reactor design appeared to be a viable approach, but incidents like the Fermi 1 disaster near Detroit cast doubt on the wisdom of using this approach. The reactor design that emerged as the dominant choice in civilian power production was the light water reactor. But light water reactors presented their own technological challenges, including most especially the risk of a massive steam explosion in the event of a power interruption to the cooling plant. In order to obviate this risk reactor designs involved multiple levels of redundancy to ensure that no such power interruption would occur. And much of the cost of construction of a modern light water power plant is dedicated to these systems -- containment vessels, redundant power supplies, etc. In spite of these design efforts, however, light water reactors at Three Mile Island and Fukushima did in fact melt down under unusual circumstances -- with particularly devastating results in Fukushima. The nuclear power industry in the United States essentially died as a result of public fears of the possibility of meltdown of nuclear reactors near populated areas -- fears that were validated by several large nuclear disasters.

What is interesting about this story is that there was an alternative reactor design that was developed by US nuclear scientists and engineers in the 1950s that involved a significantly different solution to the problem of harnessing the heat of a nuclear reaction and that posed a dramatically lower level of risk of meltdown and radioactive release. This is the molten salt reactor, first developed at the Oak Ridge National Laboratory facility in the 1950s. This was developed as part of the loopy idea of creating an atomic-powered aircraft that could remain aloft for months. This reactor design operates at atmospheric pressure, and the technological challenges of maintaining a molten salt cooling system are readily solved. The fact that there is no water involved in the cooling system means that the greatest danger in a nuclear power plant, a violent steam explosion, is eliminated entirely. Molten salt will not turn to steam, and the risk of a steam-based explosion is removed completely. Chinese nuclear energy researchers are currently developing a next generation of molten salt reactors, and there is a likelihood that they will be successful in designing a reactor system that is both more efficient in terms of cost and dramatically safer in terms of low-probability, high-cost accidents (link). This technology also has the advantage of making much more efficient use of the nuclear fuel, leaving a dramatically smaller amount of radioactive waste to dispose of.

So why did the US nuclear industry abandon the molten-salt reactor design? This seems to be a situation of lock-in by an industry and a regulatory system. Once the industry settled on the light water reactor design, it was implemented by the Nuclear Regulatory Commission in terms of the regulations and licensing requirements for new nuclear reactors. It was subsequently extremely difficult for a utility company or a private energy corporation to invest in the research and development and construction costs that would be associated with a radical change of design. There is currently an effort by an American company to develop a new-generation molten salt reactor, and the process is inhibited by the knowledge that it will take a minimum of ten years to gain certification and licensing for a possible commercial plant to be based on the new design (link).

This story illustrates the possibility that a process of technology development may get locked into a particular approach that embodies substantial public risk, and it may be all but impossible to subsequently adopt a different approach. In another context Thomas Hughes refers to this as technological momentum, and it is clear that there are commercial, institutional, and regulatory reasons for this "stickiness" of a major technology once it is designed and adopted. In the case of nuclear power the inertia associated with light water reactors is particularly unfortunate, given that it blocked other solutions that were both safer and more economical.

Section 8

SOCIAL RESILIENCE AND DISASTER

The fact of large-scale technology failure has come up fairly often in *Understanding Society*. There are a couple of reasons for this. One is that our society is highly technology-dependent, relying on more and more densely interlinked and concentrated systems of production and delivery that are subject to unexpected but damaging forms of failure. So it is a pressingly important problem for us to have a better understanding of technology failure than we do today. The other reason that examples of technology failure are frequent here is that it seems pretty clear that failures of this kind are generally social and organizational failures (in part), not simply technological failures. So the study of technology failure is a good way of examining the weaknesses and strengths of various organizational forms -- from the firm or plant to the vast regulatory agency. I have highlighted the work of Charles Perrow as being especially useful in this context, especially *Normal Accidents: Living with High-Risk Technologies* and *The Next Catastrophe: Reducing Our Vulnerabilities to Natural, Industrial, and Terrorist Disasters*.

Kathleen Tierney has studied disasters very extensively, and her recent *The Social Roots of Risk: Producing Disasters, Promoting Resilience* is an important contribution. Tierney is both an academic and a practitioner; she is an expert on earthquake science and preparedness and serves as director of the Natural Hazards Center at the University of Colorado. The topics of disaster and technology failure are linked; natural disasters (earthquakes, tsunami, hurricanes) are often the cause of ensuing technology failures of enormous magnitude. Here is Tierney's over-riding framework of analysis:

The general answer is that disasters of all types occur as a consequence of common sets of social activities and processes that are well understood on the basis of both social science theory and empirical data. Put simply, the organizing idea for this book is that disasters and their impacts are socially produced, and that the forces driving the production of disaster are embedded in the social order itself. As the case studies and research findings discussed *throughout the book will show, this is equally true whether the culprit in question is a hurricane, flood, earthquake, or a bursting speculative bubble. The origins of disaster lie not in nature, and not in technology, but rather in the ordinary everyday workings of society itself. (4-5)*

This is one of Tierney's key premises -- that disasters are socially produced and socially constituted. Her other major theme is the notion of resilience -- the idea that social characteristics exist that make one set of social arrangements more resilient than another to harm in the face of natural catastrophe. Features of resilience involve --

preexisting, planned, and naturally emerging activities that make societies and communities better able to cope, adapt, and sustain themselves when disasters occur, and also to develop ways of recovering following such events. (5)

Tierney is often drawn to the alliteration of "risk and resilience". "Risk" is the possibility of serious disturbance to the integrity of a system. "Risk" is a compound of likelihood of a type of disturbance and the damage created by that eventuality. Here is Tierney's capsule definition:

Risk is commonly conceptualized as the answer to three questions: What can go wrong? How likely is it? And what are the consequences? (11)

"Resilience", by contrast, is a feature of the system in response to such a disturbance. So the concepts of risk and resilience do not operate on the same level. A more apt opposition is ***fragility*** and resilience. (Tierney sometimes refers to brittle institutions.) Some institutional arrangements are like glass -- a sharp tap and they fall into a mound of shards. Others are more like a starfish -- able to recover form and function following even very damaging encounters with the world. Both kinds of systems are subject to risk, and the probability of a given disturbance may be the same in the two instances. The difference between them is how well they recover from the realization of risk. But the damage that results from the same disturbance is much greater in a fragile system than a resilient system. And Tierney makes a crucial point for all of us in the twenty-first century: we need to be exerting ourselves to create social systems and communities that are substantially more resilient than they currently are.

A very important example of non-resilient trends in twenty-first century life is the spread of ultra-tall buildings in global cities. There are a variety of reasons why developers and urban leaders like ultra-tall structures -- reasons that largely have to do with prestige. But Tierney points out in expert detail the degree to which these buildings are unreasonably fragile in face of disaster: they shed vast quantities of glass, they concentrate people and business in a way that invites terrorist attack, they exist in vulnerable systems of electricity and water that are crucial to their hour-to-hour functioning. A major earthquake in San Francisco has the potential to leave the buildings standing but the populations living within them stranded without light or elevators, and the emergency responders one hundred flights of stairs away from the emergencies they need to confront (63ff.).

The most fundamental and intractable source of hazard for our society that Tierney highlights is the likelihood of failure of government regulatory and safety organizations to carry out their stated missions of protecting the safety and health of the public. Like Perrow in *The Next Catastrophe*, she finds instance after instance of cases where the public's interest would be best served by a regulation or prohibition of a certain kind of risky activity (residential and commercial development in flood or earthquake zones, for example) but where powerful economic interests (corporations, local developers) have the overwhelming ability to block sensible and prudent regulations in this space. "Economic power on this scale is easily translated into political power, with important consequences for risk buildup" (91). Tierney offers the case of the Japanese nuclear industry as an example of a concentrated and powerful set of organizations that were able to succeed in creating siting decisions and safety regulations that served their interests rather than the interests of the general public.

As nuclear power emerged as a major source of energy in Japan, communities were essentially bribed into accepting nuclear plants, with the promise of jobs for young workers and support for schools and community projects; also, extensive propaganda efforts were *launched.... Then, once government and industry succeeded in getting communities to accept the presence of nuclear plants, the natural tendency was to locate multiple reactors at nuclear sites to achieve economies of scale and to avoid having to repeat costly charm offensives in large numbers of communities. (92)*

In Tierney's view, the problem of regulatory capture by the economically powerful is perhaps the largest obstacle to our ability to create a rational and prudent plan for managing risks in the future (94).

The *Social Roots of Risk* is rich in detail and deeply insightful into the sociology of risk in a large democratic corporation-centered society. The hazards she identifies concerning the failure of our institutions to devise genuinely prudent policies around foreseeable risks (earthquake, hurricane, flood, terrorism, nuclear or chemical plant malfunction, train disaster, ...) are deeply alarming. The public and our governments need to absorb these lessons and design for more resilient societies and communities, exactly as Tierney and Perrow argue.

Section 9

REGULATORY THROMBOSIS

Charles Perrow is a leading researcher on the sociology of organizations, and he is a singular expert on accidents and system failures. Several of his books are classics in their field -- *Normal Accidents: Living with High-Risk Technologies, The Next Catastrophe: Reducing Our Vulnerabilities to Natural, Industrial, and Terrorist Disasters, Organizing America: Wealth, Power, and the Origins of Corporate Capitalism*. So it is very striking to find that Perrow is highly skeptical about the ability of governmental organizations in the United States to protect the public from large failures and disasters of various kinds -- hurricanes, floods, chemical plant fires, software failures, terrorism. His assessment of organizations such as the Federal Emergency Management Agency, the Department of Homeland Security, or the Nuclear Regulatory Commission is dismal. Here is his summary assessment of the Department of Homeland Security:

> *We should not expect too much of organizations, but the DHS is extreme in its dysfunctions. As with all organizations, the DHS has been used by its masters and outsiders for purposes that are beyond its mandate, and the usage of the DHS has been extreme. One major user of the DHS is Congress. While Congress is the arm of the government that is closest to the people, it is also the one that is most influenced by corporations and local interest groups that do not have the interests of the larger community in mind. (The Next Catastrophe, kl 205)*

The most alarming chapters of *The Next Catastrophe* concern the failures of US agencies to effectively and intelligently organize preparations that will genuinely make us safer. Perrow provides extended analyses of the Department of Homeland Security, FEMA, and the Nuclear Regulatory Commission in their respective functions -- securing the country against the consequences of terrorist attack, preparing for and responding to major environmental disasters like Katrina, and securing nuclear power plants and spent fuel storage dumps against accident and attack. In chapter after chapter he documents the most egregious and frightening failures of each of these agencies.

The level of organizational ineptitude that he documents in the performance of these agencies is staggering -- one has the impression that a particularly gifted group of high school seniors could have done a better job of responding to the Katrina disaster. ("You're doing a great job, Brownie!") And the disarray that he documents in these organizations is genuinely frightening. He walks through plausible scenarios through which a group of a dozen determined attackers could disable the cooling systems for spent fuel rods at existing nuclear power plants, with catastrophic release of radiation affecting millions of people within 50 miles of the accident. (These scenarios are all the more believable now that we've seen what happened to the cooling ponds at Fukushima.)

What this all suggests is that the U.S. government and our political culture do a particularly bad job of creating organizational intelligence in response to crucial national challenges. By this I mean an effective group of bureaus with a clear mission, committed executive leadership, and consistent communication and collaboration among agencies and a demonstrated ability to formulate and carry out rational plans in addressing identified risks. (Perrow's general assessment of the French nuclear power system seems to be that it is more effective in maintaining safe operations and protecting nuclear materials against attack.) And the US government's ability to provide this kind of intelligent risk abatement seems particularly weak.

Perrow doesn't endorse the general view that organizations can never succeed in accomplishing the functions we assign to them -- hospitals, police departments, even labor unions. Instead, there seem to be particular reasons why large regulatory agencies in the United States have proven particularly inept, in his assessment. The most faulty organizations are those that are designed to regulate risky activities and those that are charged to create prudent longterm plans for the future that seem particularly suspect, in his account. So what are those reasons for failure in these kinds of organizations?

One major part of his assessment focuses on the role that economic and political power plays in deforming the operations of major organizations to serve the interests of the powerful. Regulatory agencies are "captured" by the powerful industries they are supposed to oversee, whether through influence on the executive branch or through merciless lobbying of the legislative branch. Energy companies pressure the Congress and the NRC to privatize security at nuclear power plants -- with what would otherwise be comical results when it comes to testing the resulting level of security at numerous plants. Private security forces are given advance notice of the time and nature of the simulated attack -- and even so half the attacks are successful.

Another major source of dysfunction that Perrow identifies in the case of the Department of Homeland Security is the workings of Congressional politics. Committee chairs resist losing scope for their committees, so the oversight process remains disjointed and disruptive to the functioning of the agencies. Senators from low-population states block the distribution of DHS funds to enhance the ability of first-responders to be effective in the first hours of an incident, in order to get higher levels of funding for their low-risk populations. So California receives only roughly 13% the per-capita level of funding for anti-terrorism functions that Vermont or Wyoming receive. And of course the funds available through Homeland Security become a major prize for lobbyists, corporations, and other interested parties -- with resulting congressional pressure on DHS strategies and priorities.

Another culprit in this story of failure is the conservative penchant for leaving everything to private enterprise. As Michael Brown put the point during his tenure as director of FEMA, “The general idea—that the business of government is not to provide services, but to make sure that they are provided—seems self-evident to me” (kl 1867). The sustained ideological war against government regulation that has been underway since the Reagan administration has had disastrous consequences when it comes to safety. Activities like nuclear power generation, chemical plants, air travel, drug safety, and residential development in hurricane or forest fire zones are all too risky to be left to private initiative and self-regulation. We need strong, well-resourced, well-staffed, and independent regulatory systems for these activities, and increasingly our scorecard on each of these dimensions is in the failing range.

Overall it appears that Perrow believes that agencies like DHS and FEMA would function better if they were under clear authority of the executive branch rather than Congressional oversight and direction. Presidential authority would not guarantee success -- witness George W. Bush's hapless management of the first iteration of Homeland Security within the White House -- but the odds are better. With a President with a clearly stated and implemented priority for effective management of the risk of terrorism, the planning and coordination needed would have a greater likelihood of success.

It often sounds as though Perrow is faulting these organizations for defects that are inherent in all large organizations. But it seems more fair to say that his analysis does not identify a general feature of organizations that leads to failure in these cases, but rather a situational fact having to do with the power of business to resist regulation and the susceptibility of Congress and the President to political pressures that hamstring effective regulatory organizations. Perrow does refer to specific organizational hazards -- bad executive leadership, faltering morale, inability to collaborate across agencies, excessively hierarchical architecture -- but the heart of his argument lies elsewhere. The key set of problems spiral back to the inordinate power that corporations have in the United States, and the distortions they create in Congress and the executive branch. The risks that any sober and independent assessment would identify as highest priority are ignored in pursuit of more immediate political or personal gain. It is specifics of the US political system rather than general defects of large organizations per se that lead to the bad outcomes that Perrow identifies. There are strong democracies that do a much better job of regulating risky industries and planning for disasters than we do -- for example, France and Germany. (Here is a discussion of nuclear safety systems in France published in Nature and a discussion of nuclear safety in Germany published by Nature Conservation and Nuclear Safety.)

It is significant that even though Perrow endorses the need for strengthened regulatory agencies, he doesn't think this would be enough to prevent major catastrophes in the future. So he advocates strongly for reducing the concentration of hazards and populations. As a society, he argues, we need to come to grips with the fact that there are some kinds of activities we should simply not engage in anymore -- intensive residential building in hurricane and forest fire zones, placement of chemical and nuclear plants near cities, routing rail tankers of chlorine through cities like Baltimore and Chicago. (For that matter, a reasonable conclusion one can draw from his account of near-disasters at Indian Point in New York and Davis-Besse in Toledo, is that nuclear power is simply too high a risk to continue to tolerate.) Here is a clear statement of the gravity of culture change this would require:

> *But what if FEMA were given a mandate to deal with settlement density, escape routes, building codes, and concentrations of hazardous materials in vulnerable sites? We would need a change in out mindset to make basic vulnerabilities such as the size of cities in risky areas and the amounts of hazardous materials in urban areas as high a priority as rescue and relief. (kl 1141)*

So who will provide the political will that is needed to reverse course on nuclear and chemical regulation? The public seems to believe (falsely, it would appear) that the NRC is a rigorous and independent agency and that nuclear plants are unlikely to melt down. There isn't much public concern about these risks, and legislators are therefore free to ignore them as well. (Here is an earlier post on "quiet politics" that is relevant.) So where will the political demand for strong regulation come from? Will we need to wait for the bad news we've managed by good fortune to have avoided up to this point?

Section 10

SOVIET NUCLEAR DISASTERS: KYSHTYM

The 1986 meltdown of reactor number 4 at the Chernobyl Nuclear Power Plant was the greatest nuclear disaster the world has yet seen. Less well known is the Kyshtym disaster in 1957, which resulted in a massive release of radioactive material in the Eastern Ural region of the Soviet Union. This was a catastrophic underground explosion at a nuclear storage facility near the Mayak power plant in the Eastern Ural region of the USSR. Information about the disaster was tightly restricted by Soviet authorities, with predictably bad consequences.

Zhores Medvedev was one of the first qualified scientists to provide information and hypotheses about the Kyshtym disaster. His book *Nuclear Disaster in the Urals* was written while he was in exile in Great Britain and appeared in 1980. It is fascinating to learn that his reasoning is based on his study of ecological, biological, and environmental research done by Soviet scientists between 1957 and 1980. Medvedev was able to piece together the extent of contamination and the general nature of the cause of the event from basic information about radioactive contamination in lakes and streams in the region included incidentally in scientific reports from the period.

It is very interesting to find that scientists in the United States were surprisingly skeptical about Medvedev's assertions. W. Stratton et al published a review analysis in Science in 1979 (link) that found Medvedev's reasoning unpersuasive.

A steam explosion of one tank is not inconceivable but is most improbable, because the heat generation rate from a given amount of fission products is known precisely and is predictable. Means to dissipate this heat would be a part of the design and could be made highly reliable. (423)

They offer an alternative hypothesis about any possible radioactive contamination in the Kyshtym region -- the handful of multimegaton nuclear weapons tests conducted by the USSR in the Novaya Zemlya area.

We suggest that the observed data can be satisfied by postulating localized fallout (perhaps with precipitation) from explosion of a large nuclear weapon, or even from more than one explosion, because we have no limits on the length of time that fallout continued. (425)

And they consider weather patterns during the relevant time period to argue that these tests could have been the source of radiation contamination identified by Medvedev. Novaya Zemlya is over 1000 miles north of Kyshtym (20 degrees of latitude). So the fallout from the nuclear tests may be a possible alternative hypothesis, but it is farfetched. They conclude:

We can only conclude that, though a radiation release incident may well be supported by the available evidence, the magnitude of the incident may have been grossly exaggerated, the source chosen uncritically, and the dispersal mechanism ignored. Even so we find it hard to believe that an area of this magnitude could become contaminated and the event not discussed in detail or by more than one individual for more than 20 years. (425)

The heart of their skepticism depends on an entirely indefensible assumption: that Soviet science, engineering, and management were entirely capable of designing and implementing a safe system for nuclear waste storage. They were perhaps right about the scientific and engineering capabilities of the Soviet system; but the management systems in place were woefully inadequate. Their account rested on an assumption of straightforward application of engineering knowledge to the problem; but they failed to take into account the defects of organization and oversight that were rampant within Soviet industrial systems. And in the end the core of Medvedev's claims have been validated.

Another official report was compiled by Los Alamos scientists, released in 1982, that concluded unambiguously that Medvedev was mistaken, and that the widespread ecological devastation in the region resulted from small and gradual processes of contamination rather than a massive explosion of waste materials (link). Here is the conclusion put forward by the study's authors:

What then did happen at Kyshtym? A disastrous nuclear accident that killed hundreds, injured thousands, and contaminated thousands of square miles of land? Or, a series of rela-

tively minor incidents, embellished by rumor, and severely compounded by a history of sloppy practices associated with the complex? The latter seems more highly probable.

So Medvedev is dismissed.

After the collapse of the USSR voluminous records about the Kyshtym disaster became available from secret Soviet files, and those records make it plain that US scientists badly misjudged the nature of the Kyshtym disaster. Medvedev was much closer to the truth than were Stratton and his colleagues or the authors of the Los Alamos report.

A scientific report based on Soviet-era documents that were released after the fall of the Soviet Union appeared in the Journal of Radiological Protection in 2017 (A V Akleyev et al 2017; link). Here is their brief description of the accident:

Starting in the earliest period of Mayak PA activities, large amounts of liquid high-level radioactive waste from the radiochemical facility were placed into long-term controlled storage in metal tanks installed in concrete vaults. Each full tank contained 70–80 tons of radioactive wastes, mainly in the form of nitrate compounds. The tanks were water-cooled and equipped with temperature and liquid-level measurement devices. In September 1957, as a result of a failure of the temperature-control system of tank #14, cooling-water delivery became insufficient and radioactive decay caused an increase in temperature followed by complete evaporation of the water, and the nitrate salt deposits were heated to 330 °C–350 °C. The thermal explosion of tank #14 occurred on 29 September 1957 at 4:20 pm local time. At the time of the explosion the activity of the wastes contained in the tank was about 740 PBq [5, 6]. About 90% of the total activity settled in the immediate vicinity of the explosion site (within distances less than 5 km), primarily in the form of coarse particles. The explosion gave rise to a radioactive plume which dispersed into the atmosphere. About 2 × 106 Ci (74PBq) was dispersed by the wind (north-northeast direction with wind velocity of 5–10 m s−1) and caused the radioactive trace along the path of the plume [5]. Table 1 presents the latest estimates of radionuclide composition of the release used for reconstruction of doses in the EURT area. The mixture corresponded to uranium fission products formed in a nuclear reactor after a decay time of about 1 year, with depletion in 137Cs due to a special treatment of the radioactive waste involving the extraction of 137Cs [6]. (R20-21)

Here is the region of radiation contamination (EURT) that Akleyev et al identify:

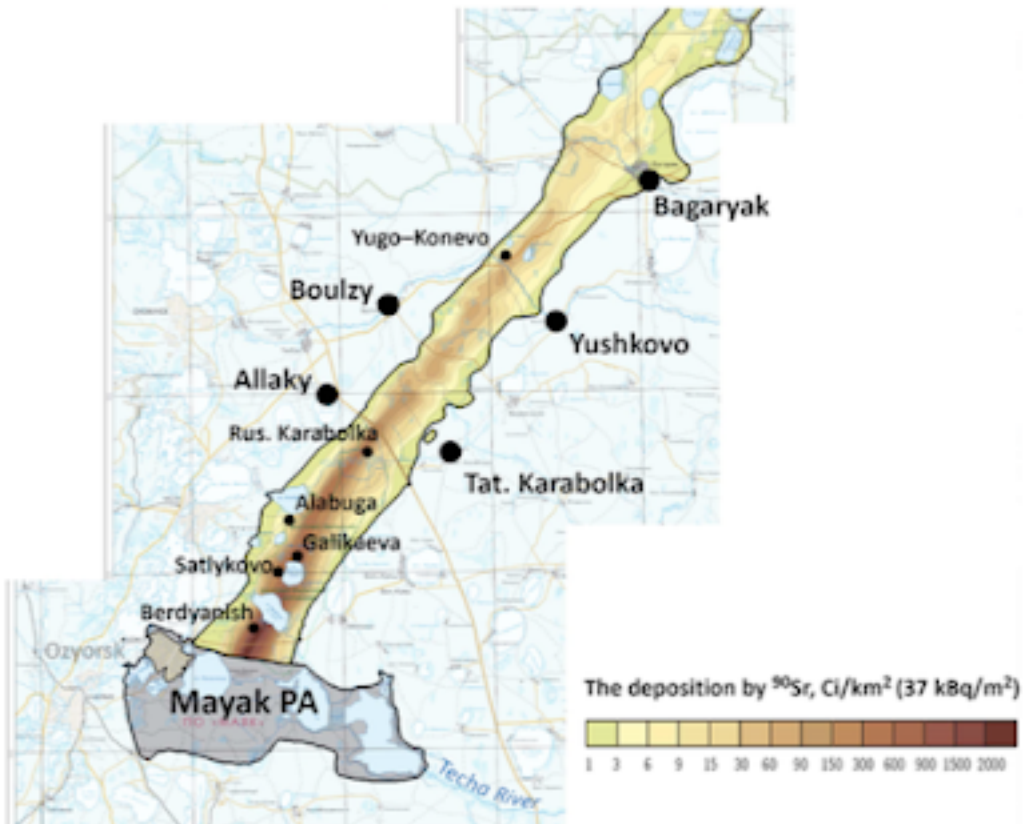


Figure 1. Schematic map of ^{90}Sr deposition in EURT territories in 1957 in Chelyabinsk Oblast (according to [8, 12]). Large circles indicate the reference non-evacuated settlements selected by URCRM for long-term monitoring; small circles indicate the evacuated settlements discussed in the present study.

This region represents a large area encompassing 23,000 square kilometers (8,880 square miles). Plainly Akleyev et al describe a massive disaster including a very large explosion in an underground nuclear waste storage facility, large-scale dispersal of nuclear materials, and evacuation of population throughout a large region. This is very close to the description provided by Medvedev.

A somewhat surprising finding of the Akleyev study is that the exposed population did not show dramatically worse health outcomes and mortality relative to unexposed populations. For example, "Leukemia mortality rates over a 30-year period after the accident did not differ from those in the group of unexposed people" (R30). Their epidemiological study for cancers overall likewise indicates only a small effect of accidental radiation exposure on cancer incidence:

The attributable risk (AR) of solid cancer incidence in the EURTC, which gives the proportion of excess cancer cases out of the sum of excess and baseline cases, calculated according to the linear model, made up 1.9% over the whole follow-up period. Therefore, only 27 cancer cases out of 1426 could be associated with accidental radiation exposure of the EURT *population. AR is highest in the highest dose groups (250–500 mGy and >500 mGy) and exceeds 17%.*

So why did the explosion occur? James Mahaffey examines the case in detail in Atomic Accidents: A History of Nuclear Meltdowns and Disasters: From the Ozark Mountains to Fukushima. Here is his account:

In the crash program to produce fissile bomb material, a great deal of plutonium was wasted in the crude separation process. Production officials decided that instead of being dumped irretrievably into the river, the plutonium that had failed to precipitate out, remaining in the extraction solution, should be saved for future processing. A big underground tank farm was built in 1953 to hold processed fission waste. Round steel tanks were installed in banks of 20, sitting on one large concrete slab poured at the bottom of an excavation, 27 feet deep. Each bank was equipped with a heat exchanger, removing the heat buildup from fission-product decay using water pipes wrapped around the tanks. The tanks were then buried under a backfill of dirt. The tanks began immediately to fill with various waste solutions from the extraction plant, with no particular distinction among the vessels. The tanks contained all the undesirable fission products, including cobalt-60, strontium-90, and cesium-137, along with unseparated plutonium and uranium, with both acetate and nitrate solutions pumped into the same volume. One tank could hold probably 100 tons of waste product.

In 1956, a cooling-water pipe broke leading to one of the tanks. It would be a lot of work to dig up the tank, find the leak, and replace the pipe, so instead of going to all that trouble, the engineers in charge just turned off the water and forgot about it.

A year passed. Not having any coolant flow and being insulated from the harsh Siberian winter by the fill dirt, the tank retained heat from the fission-product decay. Temperature inside reached 660 ° Fahrenheit, hot enough to melt lead and cast bullets. Under this condition, the nitrate solutions degraded into ammonium nitrate, or fertilizer, mixed with acetates. The water all boiled away, and what was left was enough solidified ANFO explosive to blow up Sterling Hall several times, being heated to the detonation point and laced with dangerous nuclides. [189]

Sometime before 11: 00 P.M. on Sunday, September 29, 1957, the bomb went off, throwing a column of black smoke and debris reaching a kilometer into the sky, accented with larger *fragments burning orange-red. The 160-ton concrete lid on the tank tumbled upward into the night like a badly thrown discus, and the ground thump was felt many miles away. Residents of Chelyabinsk rushed outside and looked at the lighted display to the northwest, as 20 million curies of radioactive dust spread out over everything sticking above ground. The high-level wind that night was blowing northeast, and a radioactive plume dusted the Earth in a tight line, about 300 kilometers long. This accident had not been a runaway explosion in an overworked Soviet production reactor. It was the world's first "dirty bomb," a powerful chemical explosive spreading radioactive nuclides having unusually high body burdens and guaranteed to cause havoc in the biosphere. The accidentally derived explosive in the tank was the equivalent of up to 100 tons of TNT, and there were probably 70 to 80 tons of radioactive waste thrown skyward. (KL 5295)*

So what were the primary organizational and social causes of this disaster? One is the haste created in nuclear design and construction created by Stalin's insistence on moving forward the Soviet nuclear weapons program as rapidly as possible. As is evident in the Chernobyl case as well, the political pressures on engineers and managers that followed from these political priorities often led to disastrous decisions and actions. A second is the institutionalized system of secrecy that surrounded industry generally, the military specifically, and the nuclear industry most especially. A third is the casual attitude taken by Soviet officials towards the health and wellbeing of the population. And a final cause highlighted by Mahaffey's account is the low level of attention given at the plant level to safety and maintenance of highly risky facilities. Stratton et al based their analysis on the fact that the heat-generating characteristics of nuclear waste were well understood and that effective means existed for controlling those risks. That may be, but what they failed to anticipate is that these risks would be fundamentally disregarded on the ground and in the supervisory system above the Kyshtym reactor complex.

(It is interesting to note that Mahaffey himself underestimates the amount of information that is now available about the effects of the disaster. He writes that "studies of the effects of this disaster are extremely difficult, as records do not exist, and previous residents are hard to track down" (kl 5330). But the Akleyev study mentioned above provides extensive health details about the affected population made possible as a result of data collected during Soviet times and concealed.)

Section 11

THE 737 MAX DISASTER AS AN ORGANIZATIONAL FAILURE

The topic of the organizational causes of technology failure comes up frequently in Understanding Society. The tragic crashes of two Boeing 737 MAX aircraft in the past year present an important case to study. Is this an instance of pilot error (as has occasionally been suggested)? Is it a case of engineering and design failures? Or are there important corporate and regulatory failures that created the environment in which the accidents occurred, as the public record seems to suggest?

The formal accident investigations are not yet complete, and the FAA and other air safety agencies around the world have not yet approved the aircraft for flight following the suspension of certification following the second crash. There will certainly be a detailed and expert case study of this case at some point in the future, and I will be eager to read the resulting book. In the meantime, though, it is useful to bring the perspectives of Charles Perrow, Diane Vaughan, and Andrew Hopkins to bear on what we can learn about this case from the public media sources that are available. The preliminary sketch of a case study offered below is a first effort and is intended simply to help us learn more about the social and organizational processes that govern the complex technologies upon which we depend. Many of the dysfunctions identified in the safety literature appear to have had a role in this disaster.

I have made every effort to offer an accurate summary based on publicly available sources, but readers should bear in mind that it is a preliminary effort.

The key conclusions I've been led to include these:

The updated flight control system of the aircraft (MCAS) created the conditions for crashes in rare flight conditions and instrument failures.

- Faults in the AOA sensor and the MCAS flight control system persisted through the design process
- pilot training and information about changes in the flight control system were likely inadequate to permit pilots to override the control system when necessary

There were fairly clear signs of organizational dysfunction in the development and design process for the aircraft:

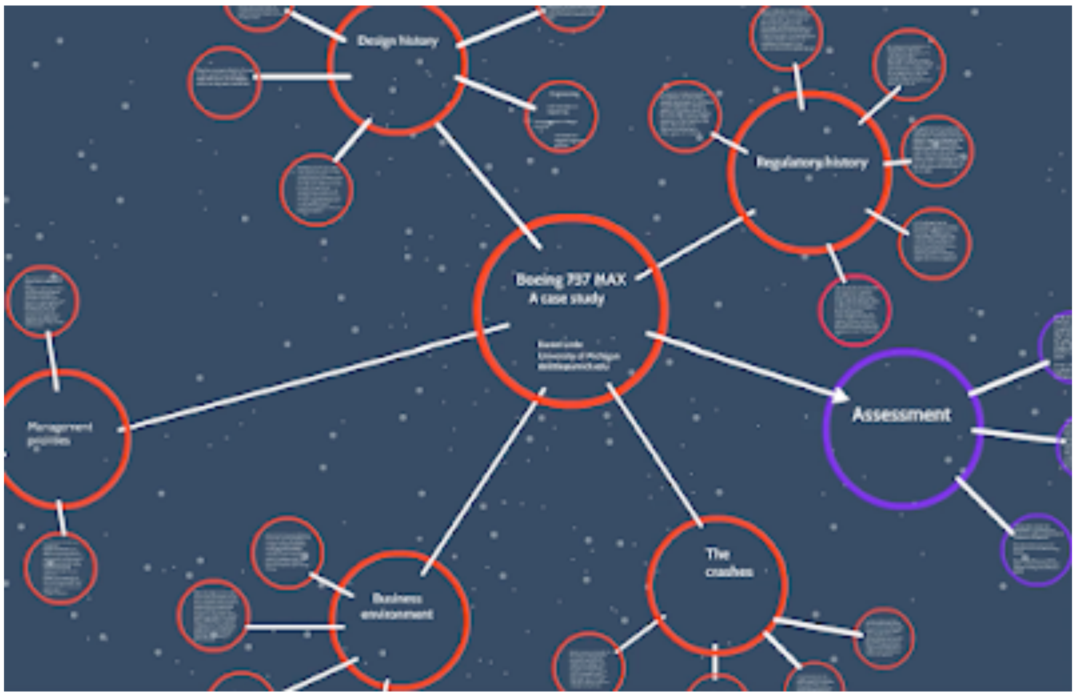
- Disempowered mid-level experts (engineers, designers, software experts)
- Inadequate organizational embodiment of safety oversight
- Business priorities placing cost savings, timeliness, profits over safety
- Executives with divided incentives
- Breakdown of internal management controls leading to faulty manufacturing processes

Cost-containment and speed trumped safety. It is hard to avoid the conclusion that the corporation put cost-cutting and speed ahead of the professional advice and judgment of the engineers. Management pushed the design and certification process aggressively, leading to implementation of a control system that could fail in foreseeable flight conditions.

The regulatory system seems to have been at fault as well, with the FAA taking a deferential attitude towards the company's assertions of expertise throughout the certification process. The regulatory process was "outsourced" to a company that already has inordinate political clout in Congress and the agencies.

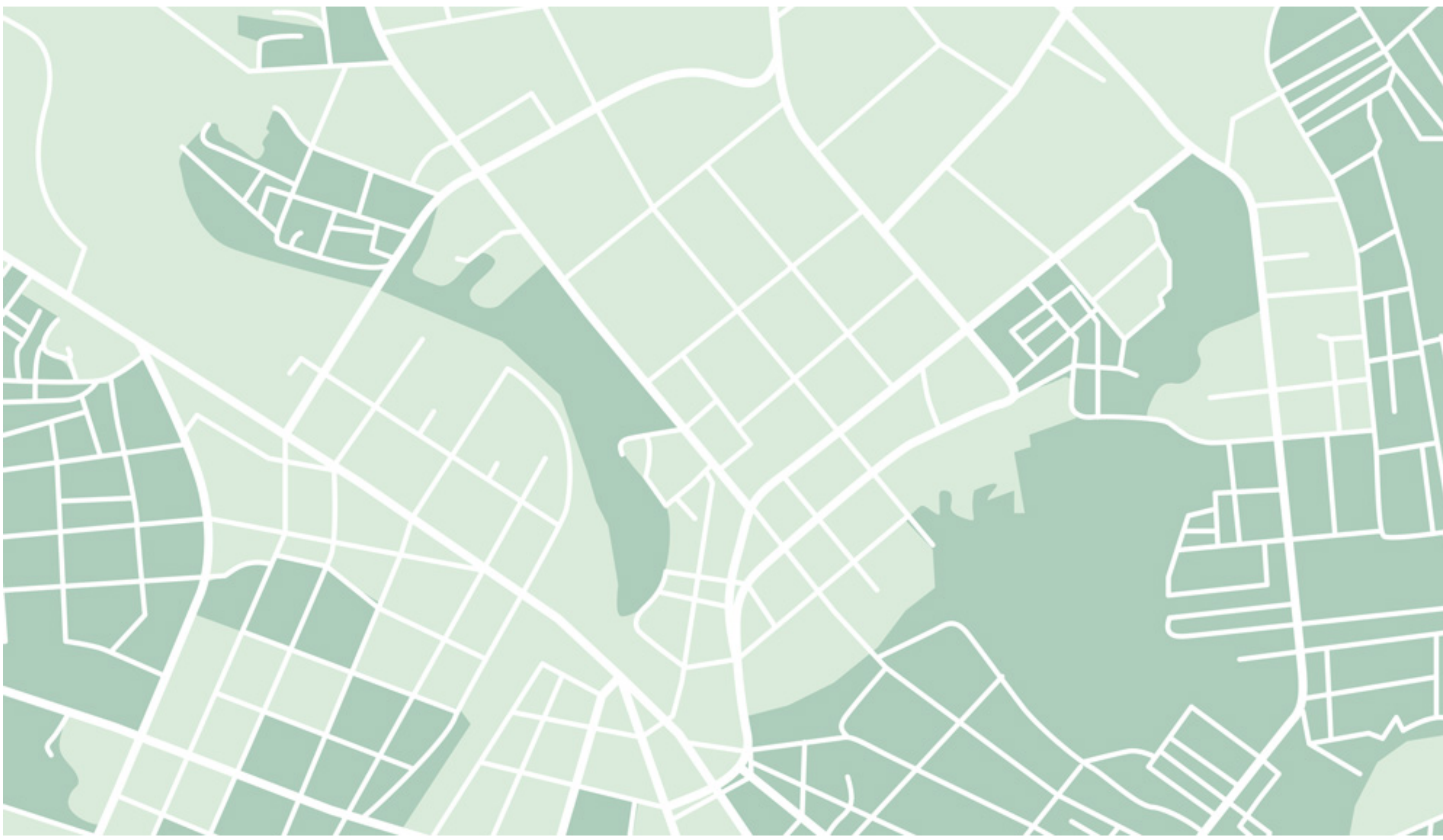
- Inadequate government regulation
- FAA lacked direct expertise and oversight sufficient to detect design failures.
- Too much influence by the company over regulators and legislators

Here is a video presentation of the case as I currently understand it (link).


Design history
Regulatory history
Boeing 737 MAX
A case study
Assessment
Management priorities
Business environment
The crashes

Chapter 6

TECHNOLOGY AND THE ENVIRON-MENT



Lorem Ipsum

1. Lorem ipsum dolor sit amet, consectetur.
2. Nulla et urna convallis nec quis blandit odio mollis.
3. Sed metus libero cing elit, lorem ipsum. Quis que euismod bibendum.

Lorem ipsum dolor sit amet, ligula suspendisse nulla pretium, rhoncus tempor placerat fermentum, enim integer ad vestibulum volutpat. Nisl rhoncus turpis est, vel elit, congue wisi enim nunc ultricies sit, magna tincidunt. Maecenas aliquam maecenas ligula nostra, accumsan taciti. Sociis mauris in integer, a dolor netus non dui aliquet, sagittis felis sodales, dolor sociis mauris, vel eu est libero cras. Interdum at. Eget habitasse elementum est, ipsum purus pede porttitor class, ut lorem adipiscing, aliquet sed auctor, imperdiet arcu per diam dapibus libero duis. Enim eros in vel, volutpat nec pellentesque leo, temporibus scelerisque nec.

Ac dolor ac adipiscing amet bibendum nullam, massa lacus molestie ut libero nec, diam et, pharetra sodales eget, feugiat ullamcorper id tempor eget id vitae. Mauris pretium eget aliquet, lectus tincidunt. Porttitor mollis imperdiet libero senectus pulvinar. Etiam molestie mauris ligula eget laoreet, vehicula eleifend. Repellat orci eget erat et, sem cum, ultricies sollicitudin amet eleifend dolor nullam erat, malesuada est leo ac. Varius natoque turpis elementum est. Duis montes, tellus lobortis lacus amet arcu et. In vitae vel, wisi at, id praesent bibendum libero faucibus porta egestas, quisque praesent ipsum fermentum placerat tempor.

Curabitur auctor, erat mollis sed fusce, turpis vivamus a dictumst congue magnis. Aliquam amet ullamcorper dignissim molestie, sed mollis. Tortor vitae tortor eros wisi facilisis.Consectetuer arcu ipsum ornare pellentesque vehicula, in vehicula diam, ornare magna erat felis wisi a risus. Justo fermentum id. Malesuada eleifend, tortor eros.

Section 1

FLOOD PLAINS AND LAND USE

An increasingly pressing consequence of climate change is the rising threat of flood in coastal and riverine communities. And yet a combination of Federal and local policies have created land use incentives that have led to increasing development in flood plains since the major floods of the 1990s and 2000s (Mississippi River 1993, Hurricane Katrina 2005, Hurricane Sandy 2016, ...), with the result that economic losses from flooding have risen sharply. Many of those costs are born by tax payers through Federal disaster relief and subsidies to the Federal flood insurance program.

Christine Klein and Sandra Zellmer provide a highly detailed and useful review of these issues in their brilliant SMU Law Review article, "Mississippi River Stories: Lessons from a Century of Unnatural Disasters" (link). These arguments are developed more fully in their 2014 book Mississippi River Tragedies: A Century of Unnatural Disaster. Klein and Zellmer believe that current flood insurance policies and disaster assistance policies at the federal level continue to support perverse incentives for developers and homeowners and need to be changed. Projects and development within 100-year flood plains need to be subject to mandatory flood insurance coverage; flood insurance policies should be rated by degree of risk; and government units should have the legal ability to prohibit development in flood plains. Here are their central recommendations for future Federal policy reform:

Substantive requirements for watershed planning and management would effectuate the Progressive Era objective underlying the original Flood Control Act of 1928: treating the river and its floodplain as an integrated unit from source to mouth, "systematically and consistently," with coordination of navigation, flood control, irrigation, hydropower, and ecosystem services. To accomplish this objective, the proposed organic act must embrace five basic principles:

(1) Adopt sustainable, ecologically resilient standards and objectives;

(2) Employ comprehensive environmental analysis of individual and cumulative effects of floodplain construction (including wetlands fill); (3) Enhance federal leadership and competency by providing the Corps with primary responsibility for flood control measures, cabined by clear standards, continuing monitoring responsibilities, and oversight through probing judicial review, and supported by a secure, non-partisan funding source; (4) Stop wetlands losses and restore damaged floodplains by re-establishing natural areas that are essential for floodwater retention; and (5) Recognize that land and water policies are inextricably linked and plan for both open space and appropriate land use in the floodplain. (1535-36)

Here is Klein and Zellmer's description of the US government's response to flood catastrophes in the 1920s:

Flood control was the most pressing issue before the Seventieth Congress, which sat from 1927 to 1929. Congressional members quickly recognized that the problems were two-fold. First, Congressman Edward Denison of Illinois criticized the absence of federal leadership: "the Federal Government has allowed the people. . . to follow their own course and build their own levees as they choose and where they choose until the action of the people of one State has thrown the waters back upon the people of another State, and vice versa." Moreover, as Congressman Robert Crosser of Ohio noted, the federal government's "levees only" policy--a "monumental blunder"--was not the right sort of federal guidance. (1482-83)

In passing the Flood Control Act of 1928, congressional members were influenced by Progressive Era objectives. Comprehensive planning and multiple-use management were hallmarks of the time. The goal was nothing less than a unified, planned society. In the early 1900s, many federal agencies, including the Bureau of Reclamation and the U.S. Geological Survey, had agreed that each river must be treated as an integrated unit from source to mouth. Rivers were to be developed "systematically and consistently," with coordination of navigation, flood control, irrigation, and hydro-power. But the Corps of Engineers refused to join the movement toward watershed planning, instead preferring to conduct river management in a piecemeal fashion for the benefit of myriad local interests. (1484)

But perverse incentives were created by Federal flood policies in the 1920s that persist to the present:

Only a few decades after the 1927 flood, the Mississippi River rose up out of its banks once again, teaching a new lesson: federal structural responses plus disaster relief payouts had incentivized ever more daring incursions into the floodplain. The floodwater evaded federal efforts to control it with engineered structures, and those same structures prevented the river from finding its natural retention areas--wetlands, oxbows, and meanders--that had previously provided safe storage for floodwater. The resulting damage to affected areas was increased by orders of magnitude. The federal response to this lesson was the adoption of a nationwide flood insurance program intended to discourage unwise floodplain development and to limit the need for disaster relief. Both lessons are detailed in this section. (1486)

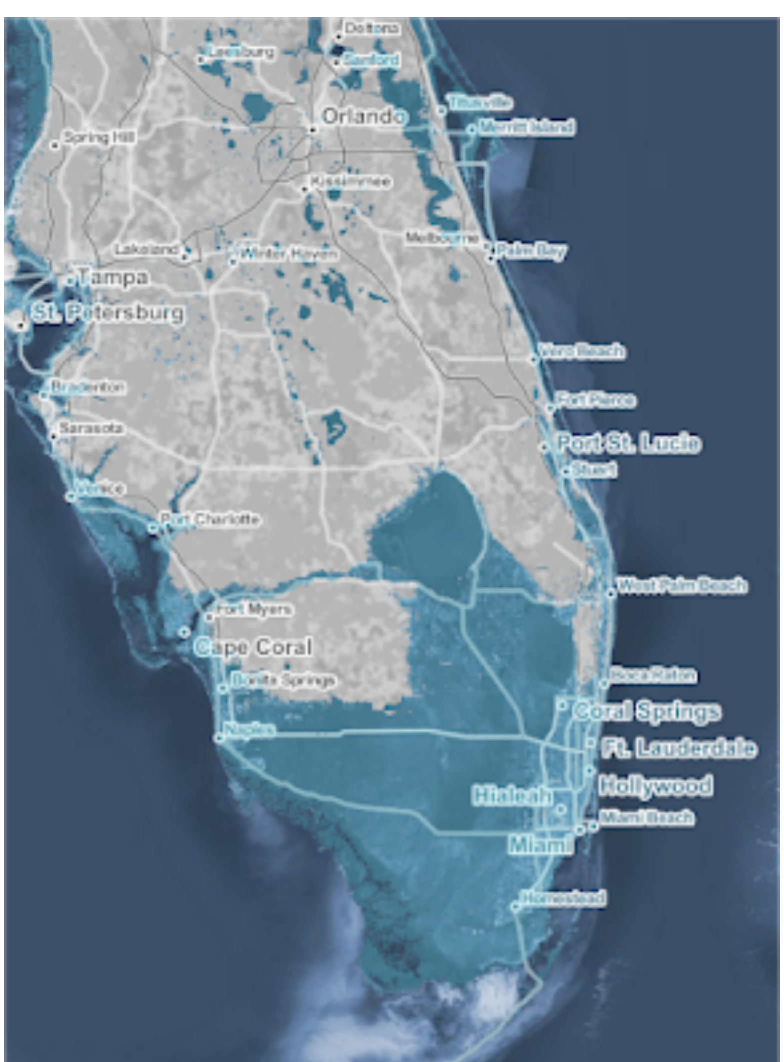
Paradoxically, navigational structures and floodplain constriction by levees, highway embankments, and development projects exacerbated the flood damage all along the rivers in 1951 and 1952. Flood-control engineering works not only enhanced the danger of floods, but actually contributed to higher flood losses. Flood losses were, in turn, used to justify more extensive control structures, creating a vicious cycle of ever-increasing flood losses and control structures. The mid-century floods demonstrated the need for additional risk-management measures. (1489)

Only five years after the program was enacted, Gilbert White's admonition was validated. Congress found that flood losses were continuing to increase due to the accelerating development of floodplains. Ironically, both federal flood control infrastructure and the availability of federal flood insurance were at fault. To address the problem, Congress passed the Flood Disaster Protection Act of 1973, which made federal assistance for construction in flood hazard areas, including loans from federally insured banks, contingent upon the purchase of flood insurance, which is only made available to participating communities. (1491)

But development and building in the floodplains of the rivers of the United States has continued and even accelerated since the 1990s.

Government policy comes into this set of disasters at several levels. First, climate policy -- the evidence has been clear for at least two decades that the human production of greenhouse gases is creating rapid climate change, including rising temperatures in atmosphere and oceans, severe storms, and rising ocean levels. A fundamental responsibility of government is to regulate and direct activities that create public harms, and the US government has failed abjectly to change the policy environment in ways that substantially reduce the production of CO_2 and other greenhouse gases. Second, as Klein and Zellmer document, the policies adopted by the US government in the early part of the twentieth century intended to prevent major flood disasters were ill conceived. The efforts by the US government and regional governments to control flooding through levees, reservoirs, dams, and other infrastructure interventions have failed, and have probably made the problems of flooding along major US rivers worse. Third, the human activities in flood plains -- residences, businesses, hotels and resorts -- have worsened the severity of the consequences of floods by elevating the cost in lives and property because of reckless development in flood zones. Governments have failed to discourage or prevent these forms of development, and the consequences have proven to be extreme (and worsening).

It is evident that storms, floods, and sea-level rise will be vastly more destructive in the decades to come. Here is a projection of the effects on the Florida coastline after a sustained period of sea-level rise resulting from a 2-degree Centigrade rise in global temperature (link):



We seem to have passed the point where it will be possible to avoid catastrophic warming. Our governments need to take strong actions now to ameliorate the severity of global warming, and to prepare us for the damage when it inevitably comes.

Section 2

POVERTY, GROWTH, AND SUSTAINABILITY

September 4, 2009

https://understandingsociety.blogspot.com/2009/09/poverty-growth-and-sustainability.html

The extent and depth of poverty in the world today is a crushing and immediate problem. The economies of most countries in the world continue to reproduce life circumstances for the extremely poor that make it all but impossible for them to participate in normal, productive lives. The Millenium development goals that were endorsed by the United Nations and other national and international organizations are still far from realization (link). At least a billion people in Asia, Africa, and Latin America live in extreme poverty (defined by the United Nations as per capita income less than a dollar a day), and a larger number are subject to hunger and malnutrition. This is an accumulation of misery and despair that cannot be tolerated. (Here's an earlier post on food security and hunger.)

So poverty alleviation must be a crucial and high priority. Nations and regions need to give their strongest efforts to enacting economic and social reforms that consistently work in the direction of reducing the scope of poverty and increasing the degree of human fulfillment that is feasible in the world. Jeffrey Sachs (*The End of Poverty: Economic Possibilities for Our Time*), Amartya Sen (), and Paul Collier (*The Bottom Billion: Why the Poorest Countries are Failing and What Can Be Done About It*) have all made good contributions on the pressing urgency of this issue and some ideas about strategies that can work.

Poverty alleviation can be understood to mean a variety of things: ensuring that families have higher disposable income, ensuring that families have access to the goods necessary to satisfy their basic needs; or creating social safety nets that help to provide the means of health, nutrition, and education that are fundamental to improving the productivity of the poor. And all of this implies two important consequences: the world economy needs to grow consistently for a very long time, and the purchasing power of the poorest 40% of the world's population needs to increase more rapidly than the growth rate. Overall inequalities need to decline as economic growth proceeds. Getting out of poverty means, among other things, having access to more of society's resources for the sake of consumption: better diet, healthcare, education, transportation, housing, clothing, and other goods. And this is critically important, because as Amartya Sen has argued throughout his career, poverty obliterates human potential and opportunity. It destroys the individual's "capabilities and functionings" for life.

Now consider the environmental side of the coin. Sustainability means designing a social and economic system that is eventually ... sustainable. It means using non-renewable resources in a way that permits future generations to have the ability to achieve the things they will need to do; it means using renewable resources in ways that permit replacement; it means managing water and air quality in such a way that we're not locked into a downward spiral of worsening quality in these essential resources; and, of course, it means managing human activity in ways that permit control of global climate change. People like James Gustave Speth (*The Bridge at the Edge of the World: Capitalism, the Environment, and Crossing from Crisis to Sustainabili-*

ty) have argued that these goals can only be met by damping down the consumption patterns of the world's affluent billion (post). But how will that work for the world's poorest billion?

Here is the hard point, however: the goals of poverty alleviation and sustainability appear on their face to be inconsistent. If we want global poverty alleviation, then it is hard to see how we can have a sustainable economy in the next fifty years; and if we insist on sustainability, then it is hard to see how we can achieve enduring and improving poverty alleviation. It appears inescapable that if we succeed in doubling the real income of the poorest 40% of the world's population, this means even greater increases in the consumption of energy and other resources. And the production of energy is the primary cause of failures of sustainability. If part of alleviating poverty in China or India means making it feasible for three times as many poor families to own a motorbike or automobile -- this implies a significant increase in the consumption of fossil fuels and the production of greenhouse gases. And yet improving mobility is an important cause and benefit of alleviating poverty. And this is true for a wide range of categories of products: being less poor means having the ability to consume more of these products.

Here is a hypothetical effort to think through the consequences of alleviating poverty for aggregate consumption of various resources from my *The Paradox of Wealth and Poverty: Mapping the Ethical Dilemmas of Global Development*.

	2000	2050	change
Population (million)	6,000	9,000	50%
GDP per capita (1998 $)	$5,000	$15,000	200%
GDP for global economy (million $)	$30,000,000	$135,000,000	350%
Minimum income	$1,000	$8,000	700%
Gini coefficient	0.700	0.500	-29%
per capita			
Energy (tons of oil)	1.70	2.75	62%
Metals	500.00	1000.00	100%
grain (kg)	250.00	500.00	100%
meat (kg)	33.00	75.00	127%
fish (kg)	50.00	75.00	50%
total output			
Energy (million tons of oil)	10,200	24,750	143%
Metals (million)	3,000,000	9,000,000	200%
Grain (million kg)	1,500,000	4,500,000	200%
Meat (million kg)	198,000	675,000	241%
Fish (million kg)	300,000	675,000	125%
agricultural output			
grain per hectare (kg)	3,000	4,500	50%
meat per hectare (kg)	700	1,000	43%
fish per hectare (kg)	800	1,200	50%

This isn't a rigorously developed model; it simply makes some basic assumptions about current and future consumption levels for the world's population and then aggregates them for the full population. It is generally recognized, for example, that there is something like a "nutrition transition" that occurs as income rises for poor people, as they have the ability to substitute more meat and fish for grains and vegetables in their diet. But this implies a disproportionate increase in the amount of farmland dedicated to livestock. And globally, this implies a reduction in forests and grasslands -- processes already underway in Central and South America -- which has major consequences for CO2 increases and species extinction. Likewise, this impressionistic projection implies a tripling of the consumption of metals and grain; more than tripling of meat consumption; and 2.4 times the level of energy consumption. So it seems that poverty alleviation all but guarantees that the demand for environment-stressing commodities will increase significantly more rapidly than population increase.

So here is the key question: are there pathways for increasing the quality of life of the world's poorest 40% that are also compatible with reaching a sustainable global environment? The forthcoming World Bank 2010 *World Development Report* focuses on climate change, and there should be some important new ideas there about how to combine poverty alleviation and sustainability. These are the hardest problems we face today, and we need some fresh ideas.

Section 3

IS INDUSTRIAL AGRICULTURE SUSTAINABLE?

November 13, 2007

https://understandingsociety.blogspot.com/2007/11/is-industrial-farming-sustainable.html

The world's food system depends largely on a farming system with post-green-revolution techniques: new seed varieties, substantial use of chemical fertilizers and pesticides, large-scale irrigation, machine-based cultivation, production for large markets, and separation of production from consumption by long distances. This system shows the highest productivity the world has ever seen, whether measured in terms of labor, land, or cost. And the system does a fairly good job of producing enough food for the world's 6 billion people.

But is this system sustainable?

Several large issues arise. First, the system is energy-intensive, so it poses significant demands on the petroleum economy. The use of petroleum and energy pervades the process: fuel for cultivation and transport, energy and inputs into the production of chemical fertilizers and pesticides, energy consumed in irrigation. So a part of the sustainability question has to do with the energy challenge the globe faces.

Second, industrial agriculture has massive environmental effects. Fertilizer and animal waste runoffs lead to groundwater and river pollution (extending into the Gulf of Mexico). Degradation and loss of topsoil is another large and longterm environmental effect with serious consequences for future agricultural productivity. And methane produced by large-scale cattle- and swine-rearing represents a measurable component of global warming. So the environmental effects of industrial agriculture are very large--once again raising the question of global sustainability.

Finally, industrial agriculture, and the integrated global commodity markets from which this system is inseparable, have large and destructive consequences for traditional agriculture and the communities built around traditional farming. The effect of NAFTA and the export of US corn to Mexico has been massive in its disruption of maize-based culture and communities in Mexico.

Three questions are central. First, is this system sustainable in the narrow sense, or will it collapse of its own burden of soil, water, and air pollution in the next 50 years? Second, is it a potential part of a larger sustainable global system of production and consumption from an environmental point of view? Or does global sustainability require radical change in agriculture? And finally, are there feasible alternative systems that would be less environmentally harmful, more sustainable, and less disruptive of agrarian communities? Are these alternatives scaleable to the needs of mass societies, large cities, and a global population of 6-8 billion? Can alternative systems achieve the productivity needed to feed the world's population?

Environmentalists, global justice activists, and food activists have argued that there are alternatives. The Fair Trade movement is trying to get first-world consumers to favor fair-trade-certified products in their consumption--giving greater security and income to third-world farmers. Organic farming advocates argue that a system of smaller farms, organic fertilizers, innovative pest control, and farming techniques more suited to the local environment would have a smaller environmental footprint. "Local food" activists support the idea of shifting consumption towards products that can be grown locally--thus reducing transport and refrigeration and giving more of a market for small farmers.

So there are alternatives in technique and policy that could result in different farm characteristics that are more favorable from the points of view of justice, sustainability, and community. The hard question is whether these alternatives could be scaled to the volume needed to feed a mass population. And this is a question that demands careful scientific analysis.

(An excellent current critique of industrial agriculture is Michael Pollan's *The Omnivore's Dilemma: A Natural History of Four Meals*.)

Section 4

DEVELOPMENT ECONOMICS IN HISTORICAL CONTEXT

September 1, 2010

https://understandingsociety.blogspot.com/2010/09/development-economics-in-historical.html

Hollis Chenery and T. N. Srinivasan published the *Handbook of Development Economics* in 1988. It was state-of-the-art in the late 1980s. It is interesting to look back at the *Handbook* twenty-two years later to see how it stands up today.

First, the contributors. The volume is a dream-team of development thinkers from the 1970s and 1980s: Amartya Sen, Arthur Lewis, Pranab Bardhan, Joseph Stiglitz, Peter Timmer, Nancy Birdsall, Paul Streeten, and Dwight Perkins, to name only a small subset of the authors. (There are 33 essays in volumes I and II.) Several currently important figures are not represented -- Arturo Escobar, Jeffrey Sachs, and Dani Rodrik, for example. Escobar's *Encountering Development: The Making and Unmaking of the Third World* appeared in 1994; Jeffrey Sachs's *The End of Poverty: Economic Possibilities for Our Time* didn't appear until 2005; and Dani Rodrik's *One Economics, Many Recipes: Globalization, Institutions, and Economic Growth* appeared in 2008. So it is certainly true that the field has moved forward with the emergence of new voices and perspectives since 1988. But it is also true that the volume represents a very deep body of knowledge about some of the dynamics and policy choices pertaining to economic development.

More important is the question of the range of perspectives on development represented in the volume. Development thinking has tended to swing from progressive to neo-liberal over the decades. Progressives have paid more attention to distribution, poverty, and social provisioning; whereas neo-liberals have focused on markets and "getting the prices right," with little appetite for redistribution, government subsidies, or serious efforts at poverty reduction. Gunnar Myrdal, Amartya Sen, and Arturo Escobar represent three generations of progressive development theorists; perhaps Peter Timmer, Malcolm Gillis, and Jeffrey Williamson fall closer to the neo-liberal end of the spectrum. I would judge that the *Handbook* does a pretty good job of finding the middle of the spectrum. Chenery's own emphasis on the importance of redistribution in development (*Redistribution with Growth*) places him closer to the progressive end, along with Pranab Bardhan, Irma Adelman, and Lance Taylor (each of whom has a contribution in the volume). The book pays attention to "alternative approaches" to economic development as well as poverty-related issues like health and nutrition. The book does a good job of combining a clear vision of the goals of economic development -- improvement of human welfare -- with technical economic analysis of growth, labor markets, and trade. And many of the authors explicitly recognize the point that development economics benefits from theoretical pluralism; the approach is not narrowly neo-classical.

Here are a few interesting observations from several contributors:

Pranab Bardhan:

Development economics as a separate branch of economics originated in a widespread perception of the limited usefulness of orthodox economics, and even though its pristine separatism has mellowed over the years it retains to this day its contrary, unruly, if somewhat flaky, image in the eyes of mainstream economics. Standard neoclassical economics is mainly on the defensive in this terain and a number of alternative approaches clash and contend for our attention. (40)

Joseph Stiglitz:

The central questions facing development economics are: Why is it that some countries are so much poorer than others? What can be done to make them grow faster? Faster growth is needed if the gap in living standards is not to be widened even further. (94)

J.G. Williamson:

What explains the timing and the extent of the transition from a traditional rural to a modern urban society? Why does city growth speed up in early development and slow down in later stages? What role does migration play in the process, and do migrants make rational location decisions? Do urban labor markets serve to absorb urban immigrants quickly? Are rural emigrants driven by "push" conditions in the countryside or by "pull" conditions in the cities? Is the Third World "overurbanized"? (425)

T.P. Schultz:

The record of sustained modern growth in real per capita income cannot be accounted for by the accumulation of conventional units of physical capital or by the increased application of hours of labor per capita. The source of modern economic growth are sought instead in the changing quality of labor and capital, in the more comprehensive accounting of other inputs, and in change of organization, policy environment, or technology. ... Research on various aspects of the microeconomic relationshipb etween education and development has expanded rapidly, forging a consensus on questions for study and appropriate methodologies to address these questions. ... Studies across persons, households, farms, and firms have documented, first generally in the United States and then in many low income countries, strong empirical regularities between educational attainment of populations and their productivity and performance in both market and nonmarket (home) production activities. (544)

Jere Behrman and Anil Deolalikar:

Health and nutrition are important as ends in themselves and often are emphasized as critical components of basic needs in developing countries. In addition they may be channels through which productivity and distributional goals of developing societies may be pursued effectively if, as is often hypothesized, the productivity of low-income persons in work and in human- capital formation is positively affected by health and nutrition status. (633)

Several things are noteworthy in reviewing the contents and methods of the *Handbook* -- issues and perspectives that would now be regarded as crucial.

A phrase that does not occur in the volume is "Washington Consensus." This concept became current in the 1990s after being introduced by John Williamson in 1990 (link). Here is how Williamson puts his point: "The paper identifies and discusses 10 policy instruments about whose proper deployment Washington can muster a reasonable degree of consensus." He identifies ten policy goals as constituting the Washington Consensus: Fiscal Deficits, Public Expenditure Priorities, Tax Reform, Interest Rates, The Exchange Rate, Trade Policy, Foreign Direct Investment, Privatization, Deregulation, and Property Rights. It is apparent that this list is heavily tilted towards the neo-liberal end of the spectrum. By contrast, consider the Millenium Goals adopted by the United Nations in 2000 (link): End Hunger, Universal Education, Gender Equity, Child Health, Maternal Health, Combat HIV/AIDS, Environmental Sustainability, Global Partnership. The Millenium Goals are focused on ending world poverty, while the Washington Consensus is focused on achieving effective market institutions and trading systems globally. The *Handbook* isn't a sourcebook or a polemic in support of the neo-liberal agenda; but neither is it emphatic in its treatment of poverty.

Another term that does not occur in the volume is "globalization." There are discussions of international trade, migration, capital flows, transnational corporations, and credit markets -- important components of contemporary debates about globalization. But the concept space involved in the idea of economic development had not yet fully highlighted the importance of global interconnectivity.

Third, the *Handbook* gives virtually no attention to sustainability, resource depletion, and the environment. These are now regarded as crucial aspects of the challenge of economic development. Taxation, trade, and governance come in for repeated treatments; but environmental sustainability is not raised as a significant issue.

Finally, the *Handbook* doesn't give central priority to the issues of poverty alleviation and inequality that were already becoming central for some development economists, including Amartya Sen. Sen's central ideas of functionings, freedom, and capabilities are expressed in his opening chapter to the volume. But the bulk of the contributions to the Handbook don't begin with poverty, but rather more specific questions about growth, modernization, trade, and population. The conceptual shift that Sen's writings would eventually bring to the field had not yet had full effect.

It is also interesting to examine the first and second editions of an important textbook on development economics that was roughly contemporary to the Handbook. Malcolm Gillis, Dwight Perkins, Michael Roemer, and Donald Snodgrass's *Economics of Development* was published a few years earlier than the *Handbook* in 1983 and 1987. There is a high degree of conceptual and organizational similarity between the two treatments of the economics of development, including topics, approaches, and models and methods.

To get a sense for how the discipline of development economics has shifted since 1988, take a look at the topics and readings included in the course on Economic Development offered by Rohini Pande and Dani Rodrik (link). Addressing poverty is the central focus in this conceptualization of the field; there is lots of attention to the components of human wellbeing (health, education, nutrition); and the syllabus pays a good deal of attention to issues of institutions and governance within the development process.

Section 5

RETREAT OF THE ELEPHANTS

April 23, 2008

https://understandingsociety.blogspot.com/2008/04/retreat-of-elephants.html

Mark Elvin's title, *The Retreat of the Elephants: An Environmental History of China*, is brilliantly chosen to epitomize his subject: the human causes of longterm environmental change in China over a four-thousand year period of history. How many of us would have guessed that elephants once ranged across almost all of China, as far to the northeast as what is now Beijing? And what was the cause of this great retreat? It was the relentless spread of agriculture and human settlement.

In other words, human activity changed the physical environment of China in such a profound way as to refigure the range and habitat of the elephant. "Chinese farmers and elephants do not mix." This story provides an expressive metaphor for the larger interpretation of environmental history that Elvin offers: that environmental history is as much a subject of social history as it is a chronology of physical and natural changes. Human beings transform their environments -- often profoundly and at great cost.

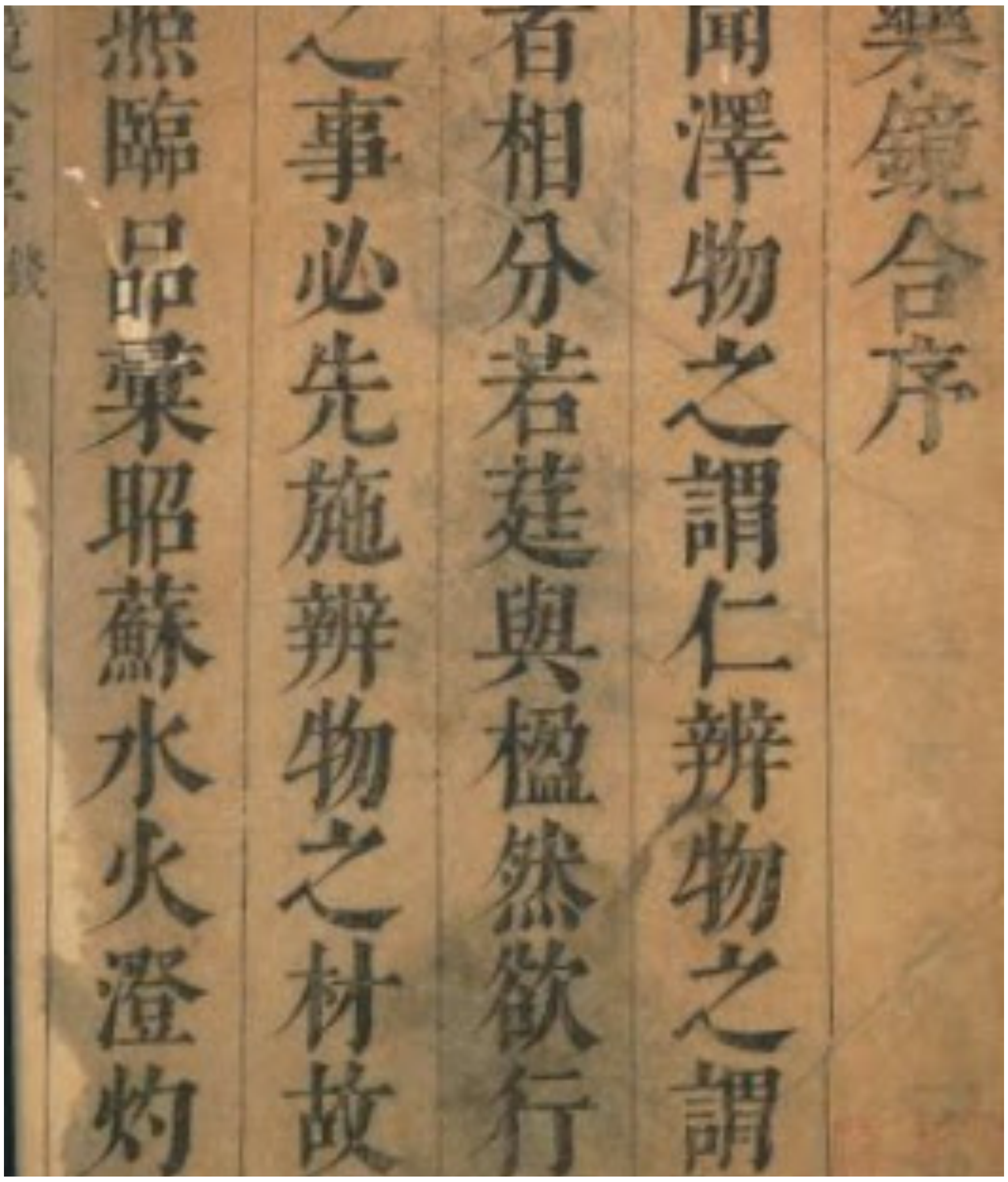
This is now a familiar story, when we consider the anthropogenic influences on global warming in the past fifty years. What Elvin's book demonstrates is that human activity is an integral part of the story in the long sweep of history as well. Nowhere is this fact more evident than in Elvin's treatment of the perennial problem of water management in China. Seawalls, canals, dikes, drainage, irrigation, desalinization, and reservoirs were all a part of China's centuries-long efforts at water control. And each of these measures had effects that refigured the next period in the water system -- the course of a river, the degree of silting of a harbor, the diminishment of a lake as a result of encroachment. (Peter Perdue's *Exhausting the Earth: State and Peasant in Hunan* tells a similar story about the fortunes of Hunan's Dongting Lake.) The waterscape of late Imperial China was very much a moving picture as human activity, deliberate policy interventions, technology innovations, and hydrology and climate interacted. There is a particular drama in seeing a centuries-long history of magistrates attempting to control the hydrology of the great rivers and deltas of the Yangtze and Yellow Rivers, to counteract silting and flooding and the massive problems that these processes entailed. Here the local officials made their best efforts to absorb the history of past interventions and their effects in order to design new systems that would obviate silting and flooding. This required planning and scientific-technical reasoning (137); it required large financial resources; and, most importantly, it required the mobilization of vast amounts of human labor to build dikes and polders. But always, in the end, the water prevailed.

Elvin's history is fascinating in a number of ways. He is an innovative writer of history, bringing new materials and new topics into Chinese historical research. His interweaving of agriculture, population growth, technology, and environmental change is masterful. He combines economic history, cultural history, and natural history in ways that bring continual new flashes of insight. He makes innovative use of literature and poetry to try to get some inklings into the attitudes and values that Chinese people brought to the environment. And he returns frequently to the dialectic of population growth and resource use -- a rising tempo of change that imposes more and more pressure on the natural environment.

(See *The High-Level Equilibrium Trap* for a discussion of one of Elvin's earlier and highly influential ideas -- the idea that Chinese agriculture had reached a stage of development by the late imperial period in which technique had been refined to the maximum possible within traditional technologies, and population had increased to the point where the agricultural system was only marginally able to feed the population. This is what he refers to as a "high-level equilibrium trap." He returns to something rather similar to this idea in Retreat of the Elephants by offering a theory of environmental exhaustion ("Concluding Remarks"): a measure of the degree to which population increase and economic growth have placed greater and greater pressure on non-renewable resources.)

Section 6

TECHNOLOGY INNOVATION IN CHINESE AGRICULTURE



October 6, 2009

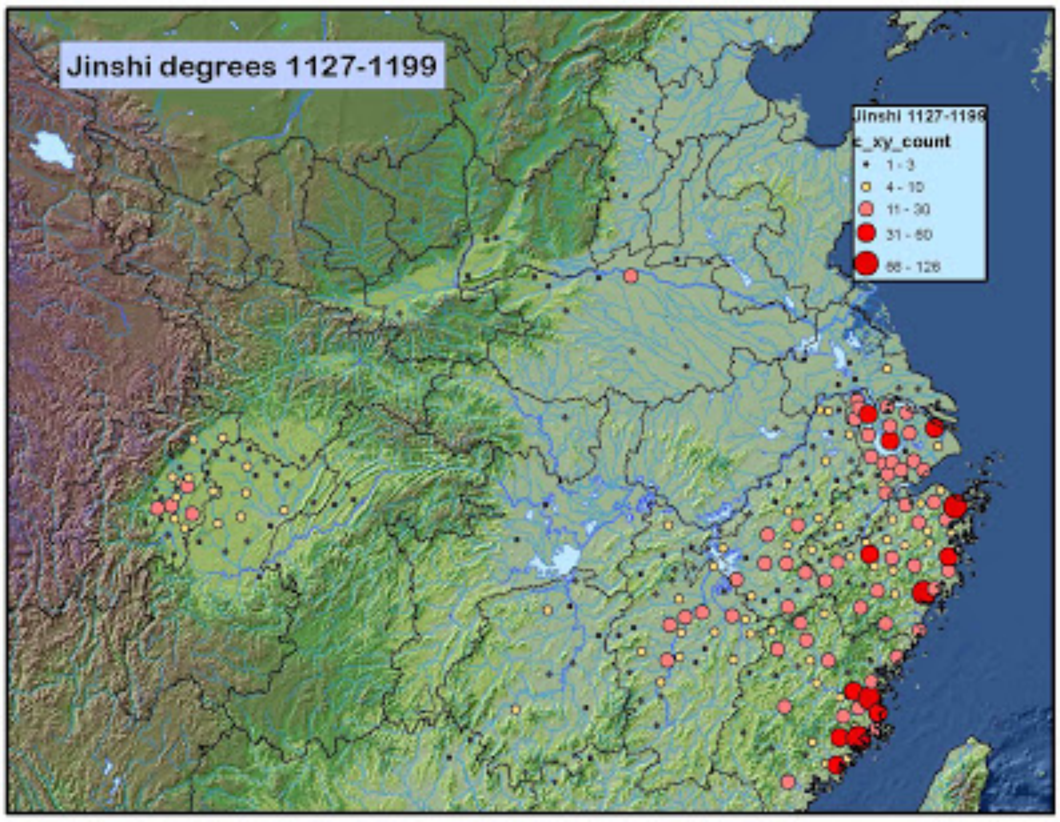
https://understandingsociety.blogspot.com/2009/10/technology-innovation-in-chinese.html

It is a commonplace in world history to observe that China had achieved a high level of sophistication in science, medicine, and astronomy by the Middle Ages, but that some unknown feature of social organization or culture blocked the further development of this science into the expansion of technology in the early modern period. Chinese culture was "blocked" from making significant technological advances during the late Ming and early Qing periods -- in spite of its scientific advantage over the West in medieval times; or so it is believed in a standard version of Chinese economic history.

A variety of hypotheses have been offered to account for this supposed fact. For example, Mark Elvin argues that China's social and demographic system created conditions for a "high-level equilibrium trap" in the early modern period in *The Pattern of the Chinese Past*. According to Elvin, Chinese social arrangements favored population growth; innovative and resourceful farmers discovered all feasible refinements of traditional agricultural techniques to refine a highly labor-intensive system of agriculture; and population expanded to the point where the whole population was at roughly the subsistence level while consuming virtually the whole of the agricultural product. There was consequently no social surplus that might have been used to invest in discovery of major innovations in agricultural technology; so the civilization was trapped. (Here is a more developed discussion of Elvin's argument.)

Other historians have speculated about potential features of Confucian culture that might have blocked the transition from scientific knowledge to technology applications. The leading Western expert on Chinese science is Joseph Needham (1900-1995), whose multi-volume studies on Chinese science set the standard in this area (*Science and Civilisation in China. Volume 1: Introductory Orientations; Clerks and Craftsmen in China and the West*). And Needham attributes China's failure to continue to make scientific progress to features of its traditional culture.

But here is a more fundamental question: is the received wisdom in fact true? Was Chinese technology unusually stagnant during the early-modern period (late Ming, early Qing)? Agriculture is a particularly important aspect of traditional economic life; so we might reformulate our question a bit more specifically: what was the status of agricultural technology in the seventeenth and eighteenth centuries (late Ming, early Qing)? (See an earlier posting on Chinese agricultural history for more on this subject.)



Economic historian Bozhong Li considers this question with respect to the agriculture of the lower Yangzi Delta in *Agricultural Development in Jiangnan, 1620-1850*. And since this was the most important agricultural region in China for centuries, his findings are important. (It was also the major cultural center of China; see the concentration of literati in the map above.) Li makes an important point about technological innovation by distinguishing between invention and dissemination. An important innovation may be discovered in one time period but only adopted and disseminated over a wide territory much later. And the economic effects of the innovation only take hold when there is broad dissemination. This was true for Chinese agriculture during the Ming period, according to Li:

The revolutionary advance in Jiangnan rice agriculture technology appeared in the late Tang and led to the emergence and development of intensive agriculture composed of double-cropping rice and wheat. But this kind of intensive agriculture in pre-Ming times was largely limited to the high-fields of western Jiangnan. In the Ming this pattern developed into what Kitada has called the 'new double-cropping system' and spread throughout Jiangnan, but only in the late Ming did it become a leading crop regime. Similar were the development and spread of mulberry and cotton farming technologies, though they were limited to particular areas and cotton technology's advances came later because cotton was introduced later. Each had its major advances in the Ming. Therefore, technology advances in Ming Jiangnan agri-

culture were certainly not inferior to those of Song times which are looked at as a period of 'farming revolution'. (40)

Li also finds that there was a significant increase in the number of crop varieties in the early Qing -- another indication of technological development. He observes, "The later the date, the greater the number of varieties. For example, in the two prefectures of Suzhou and Changzhou, 46 varieties were found in the Song, but the number rose to 118 in the Ming and 259 in the Qing" (40). And this proliferation of varieties permitted farmers to adjust their crop to local soil, water, and climate conditions -- thus increasing the output of the crop per unit of land. Moreover, formal knowledge of the properties of the main varieties increased from Ming to Qing periods; "By the mid-Qing, the concept of 'early' rice had become clear and exact, and knowledge of 'intermediate' and 'late' strains had also deepened" (42). This knowledge is important, because it indicates an ability to codify the match between the variety to the local farming environment.

Another important process of technology change in agriculture had to do with fertilizer use. Here again Li finds that there was significant enhancement, discovery, and dissemination of new uses of fertilizer in the Ming-Qing period.

A great advance in fertilizer use took place in Jiangnan during the early and mid-Qing, an advance so significant that it can be called a 'fertilizer revolution'. The advance included three aspects: (a) an improvement in fertilizer application techniques, centring on the use of top dressing; (b) progress in the processing of traditional fertilizer; and (c) an introduction of a new kind of fertilizer, oilcake. Although all three advances began to appear in the Ming, they were not widespread until the Qing. (46)

And the discovery of oilcake was very important to the increases in land productivity that Qing agriculture witnessed -- thus permitting a constant or slightly rising standard of living during a period of some population increase.

There were also advances in the use of water resources. Raising fish in ponds, for example, became an important farming activity in the late Ming period, and pond fish became a widely commercialized product in the Qing. Li describes large-scale fishing operations in Lake Tai in Jiangnan using large fishing boats with six masts to catch and transport the fish (62).

So Li's estimate of agricultural technology during the Ming period is that it was not stagnant; rather, there was significant diffusion of new crops, rotation systems, and fertilizers that led to significant increases in agricultural product during the period. "In sum, in the Jiangnan plain, land and water resources were used more rationally and fully in the early and mid-Qing than they had been in the late Ming" (64).

Two points emerge from this discussion. First, Li's account does in fact succeed in documenting a variety of knowledge-based changes in agricultural practices and techniques that led to rising productivity during the Ming-Qing period in Jiangnan. So the stereotype of "stagnant Chinese technology" does not serve us well. Second, though, what Li does not find is what we might call "science-based" technology change: for example, the discovery of chemical fertilizer, controlled experiments in rice breeding, or the use of machinery in irrigation. The innovations that he describes appear to be a combination of local adaptation and diffusion of discoveries across a broad territory.

So perhaps the question posed at the start still remains: what stood in the way of development of empirical sciences like chemistry or mechanics that would have supported science-based technological innovations in the early modern period in China?

Chapter 7

NEW ISSUES

Section 1

UNDERSTANDING AI AGENTS

Paul Churchland made a large splash in the philosophy of mind and cognitive science several decades ago when he cast doubt on the categories of "folk psychology" -- the ordinary and commonsensical concepts we use to describe and understand each other's mental lives. In Paul Churchland and Patricia Churchland, *On the Contrary: Critical Essays, 1987-1997*, Paul Churchland writes:

> *"Folk psychology" denotes the prescientific, commonsense conceptual framework that all normally socialized humans deploy in order to comprehend, predict, explain, and manipulate the behavior of . humans and the higher animals. This framework includes concepts such as belief, desire, pain pleasure, love, hate, joy, fear, suspicion, memory, recognition, anger, sympathy, intention, and so forth.... Considered as a whole, it constitutes our conception of what a person is. (3)*

Churchland does not doubt that we ordinary human beings make use of these concepts in everyday life, and that we could not dispense with them. But he is not convinced that they have a scientifically useful role to play in scientific psychology or cognitive science.

In our ordinary dealings with other human beings it is both important and plausible that the framework of folk psychology is approximately true. Our fellow human beings really do have beliefs, desires, fears, and other mental capacities, and these capacities are in fact the correct explanation of their behavior. How these capacities are realized in the central nervous system is largely unknown, though as materialists we are committed to the belief that there are such underlying neurological functionings. But eliminative materialism doesn't have a lot of credibility, and the treatment of mental states as epiphenoma to the neurological machinery isn't convincing either.

These issues had the effect of creating a great deal of discussion in the philosophy of psychology since the 1980s (link). But the topic seems all the more interesting now that tens of millions of people are interacting with Alexa, Siri, and the Google Assistant, and are often led to treat the voice as emanating from an intelligent (if not very intelligent) entity. I presume that it is clear that Alexa and her counterparts are currently "question bots" with fairly simple algorithms underlying their capabilities. But how will we think about the AI agent when the algorithms are not simple; when the agents can sustain lengthy conversations; and when the interactions give the appearance of novelty and creativity?

It turns out that this is a topic that AI researchers have thought about quite a bit. Here is the abstract of "Understanding Socially Intelligent Agents—A Multilayered Phenomenon", a fascinating 2001 article in IEEE by Perrson, Laaksolahti, and Lonnqvist (link):

The ultimate purpose with socially intelligent agent (SIA) technology is not to simulate social intelligence per se, but to let an agent give an impression of social intelligence. Such user-centred SIA technology, must consider the everyday knowledge and expectations by which users make sense of real, fictive, or artificial social beings. This folk-theoretical understanding of other social beings involves several, rather independent levels such as expectations on behavior, expectations on primitive psychology, models of folk-psychology, understanding of traits, social roles, and empathy. The framework presented here allows one to analyze and reconstruct users' understanding of existing and future SIAs, as well as specifying the levels SIA technology models in order to achieve an impression of social intelligence.

The emphasis here is clearly on the semblance of intelligence in interaction with the AI agent, not the construction of a genuinely intelligent system capable of intentionality and desire. Early in the article they write:

As agents get more complex, they will land in the twilight zone between mechanistic and living, between dead objects and live beings. In their understanding of the system, users will be tempted to employ an intentional stance, rather than a mechanistic one.. Computer scientists may choose system designs that encourage or discourage such anthropomorphism. Irrespective of which, we need to understand how and under what conditions it works.

But the key point here is that the authors favor an approach in which the user is strongly led to apply the concepts of folk psychology to the AI agent; and yet in which the underlying mechanisms generating the AI's behavior completely invalidate the application of these concepts. (This approach brings to mind Searle's Chinese room example concerning "intelligent" behavior; link.) This is clearly the approach taken by current designs of AI agents like Siri; the design of the program emphasizes ordinary language interaction in ways that lead the user to interact with the agent as an intentional "person".

The authors directly confront the likelihood of "folk-psychology" interactions elicited in users by the behavior of AI agents:

When people are trying to understand the behaviors of others, they often use the framework of folk-psychology. Moreover, people expect others to act according to it. If a person's behavior blatantly falls out of this framework, the person would probably be judged "other" in some, e.g., children, "crazies," "psychopaths," and "foreigners." In order for SIAs to appear socially intelligent, it is important that their behavior is understandable in term of the folk-psychological framework. People will project these expectations on SIA technology and will try to attribute mental states and processes according to it. (354)

And the authors make reference to several AI constructs that are specifically designed to elicit a folk-psychological response from the users:

In all of these cases, the autonomous agents have some model of the world, mind, emotions, and of their present internal state. This does not mean that users automatically infer the "correct" mental state of the agent or attribute the same emotion that the system wants to convey. However, with these background models regulating the agent's behavior the system will support and encourage the user to employ her faculty of folk-psychology reasoning onto the agent. Hopefully, the models generate consistently enough behavior to make folk-psychology a framework within which to understand and act upon the interactive characters. (355)

The authors emphasize the instrumentalism of their recommended approach to SIA capacities from beginning to end:

In order to develop believable SIAs we do not have to know how beliefs-desires and intentions actually relate to each other in the real minds of real people. If we want to create the impression of an artificial social agent driven by beliefs and desires, it is enough to draw on investigations on how people in different cultures develop and use theories of mind to understand the behaviors of others. SIAs need to model the folk-theory reasoning, not the real thing. To a shallow AI approach, a model of mind based on folk-psychology is as valid as one based on cognitive theory. (349)

This way of approaching the design of AI agents suggests that the "folk psychology" interpretation of Alexa's more capable successors will be fundamentally wrong. The agent will not be conscious, intentional, or mental; but it will behave in ways that make it almost impossible not to fall into the trap of anthropomorphism. And this in turn brings us back to Churchland and the critique of folk psychology in the human-human cases. If computer-assisted AI agents can be completely persuasive as mentally structured actors, then why are we so confident that this is not the case for fellow humans as well?

Section 2

INFORMATION TECHNOLOGY AND NEW HUMAN CAPABILITIES

January 3, 2012

https://understandingsociety.blogspot.com/2012/01/information-technology-and-new-human.html

For a billion or so of us on planet earth, we are immersed in a sea of ever-changing technology. How does technology shape us? And how do we shape technology? How do current technologies change our capacities as human beings? In what ways are we better able to fulfill our plans of life using the technologies available to us?

It goes without saying that a wealth of existing technology systems are the foundation of our current life circumstances. Electricity, commercial agriculture, large-scale logistics systems, water purification, long-distance transportation, and advanced manufacturing are critical for the lives of two-thirds of planet earth's population. And if we try to imagine what life would be like without these systems we have to go back to the lived environment of roughly 1400 in the West and perhaps 1000 in East Asia. Small population, short longevity, high maternal and infant mortality, frequent epidemic disease, grueling daily labor, and limited literacy are the baseline created by traditional agriculture and handicraft manufacture. If this is the point of comparison, then it's hard to deny that technology has improved human wellbeing.

But let's look more closely at the most recent tech revolution that we are currently experiencing, the digital information revolution. Here I'm thinking of the World Wide Web, ubiquitous web access, cheap computing power, email, jumbo databases, social media tools, and cheap global voice and video communication.

How did this new suite of technologies suddenly sweep over us? The technical side of the history is pretty well understood. The PC revolution was basically a straightforward commercialization and incremental development of computer technologies of the 1950s and 1960s. The big challenges were miniaturization and improvement of the human interface -- in other words, innovations that would permit creation of a mass market for the new devices. The personal software industry deserves its own separate mention. Of course software needed personalization -- CPM, Electric-Pencil, WordPerfect, MSDOS, Windows, Macintosh operating system. There were early innovators, and often enough those companies failed quickly. And there were a few large companies that eventually dominated. Second, the development of the first point-to-point networks permitting communication between sites was a substantial and genuine innovation. This technology would unfold into a full gauge "world-wide web" in only 15 years or so. Third, search technologies were crucial for accessing and using the millions of pages of information accessible on the web. Search tools, including especially Google, suddenly made organizing and finding information quickly a very easy, non-technical process. And a few companies jumped into the lead. The most recent wave of innovation has taken advantage of the web itself -- social networking, search, gaming, and e-commerce -- to attract users who will interact digitally through photos, video, and messaging.

It would be foolish to imagine that this technology is fundamentally different from any earlier stage of technology in its path-dependence on specific interests in society. So what were the interests that drove these developments?

Some of these shaping interests were directly related to the needs of the military. Command and control of bomber and ICBM detection systems required real time communications networks on a national scale. DARPANET was one of the early developments of these interests. Another obvious set of interests were commercial. The emerging PC technology created opportunities for large commercial success, for the entrepreneur who captured the moment. Companies like Exidy, Commodore, and Radio Shack made their efforts. But for a couple of fairly contingent reasons IBM and Apple were the big winners. And, of course, the emergence of a mass market of consumers who would be interested in buying and using these devices was critical. It is hard to imagine personal computing developing as a major industry in the former German Democratic Republic.

So it is plain that the suite of technologies that brought us the information revolution were strongly affected by governmental and commercial interests. It is also indisputable that no one could have predicted the ways these technologies would develop and interact from the starting point of 1980.

How we got here is one large question. But even more important is how this ensemble of technologies has changed us.

The positives are enormous. There is basically no limit on the range of knowledge and learning that is possible through the web. So the information revolution has offered a huge amplifier for knowledge acquisition for all of us. The fact of easily accessible information and analysis is an enhancement of our ability to understand the world.

Global communication technology is a second huge enhancement for our ability to interact with people all over the world. Scholars can collaborate in real time thanks to Skype video conferencing. Activists can interact through the same technology. Religious communities can communicate, share ideas, and disagree with each other, from Nigeria to Sao Paolo to Los Angeles.

Social networks add a third new capacity -- to create new connections with people with similar concerns and interests with whom productive interaction is possible. Twitter, Facebook, and Wordpress create micro-digital neighborhoods in which people can form surprisingly natural connections. A philosopher in Michigan becomes acquainted with a journalist in Bangkok, a mathematician in Athens, a sociology graduate in the Philippines, and a philosopher with very similar interests in Taiwan -- these are intellectual relationships that could not have occurred in a pre-web world.

So the digital revolution certainly extends human capacity and reach. But there is a negative side too. Some observers fear that the digital generation is substituting Facebook for face-to-face relationships. Skeptics argue that the so-called twitter revolutions in the Middle East can't depend on the weak bonds created by a Facebook page, and that real solidarity must proceed from more direct connections. There is real concern that hate groups can amplify their ability to mobilize through the web. Addiction to World of Warcraft and other online gaming communities seems like a real phenomenon for a significant number of people. And maybe short-form thinking (blog entries, Facebook updates, tweets) is insidiously undermining our ability to think long, coherent thoughts. So it is hard to say whether the Internet is on balance a force for extending human capabilities and social wellbeing.

The real impact of the digital revolution on the nature of human social life probably can't yet be assessed. Manuel Castells is trying to begin this process with his multi-volume *The Rise of the Network Society: The Information Age: Economy, Society, and Culture Volume I (Information Age Series)* on the Information Age. Yevgeny Morozov offers doubts about the supposedly progressive nature of the Internet in *The Net Delusion: The Dark Side of Internet Freedom*. And Sherry Turkle is exploring the personal and subjective effects of new technologies on all of us in books like *Alone Together:Why We Expect More from Technology and Less from Each Other* and *Life on the Screen: Identity in the Age of the Internet*. But realistically, we are only at the beginning of understanding the social and personal consequences of the new information and network tools that are now ubiquitous.