

Tom R. Burns and Nora Machado

The paper¹ introduces and applies actor-system dynamics (ASD), a general systems theory, to the analysis of the risks and accidents of complex, hazardous technologies and socio-technical systems. Section 1 introduces ASD theory. Section 2 applies the theory to the analysis of hazardous technologies and socio-technical systems, exposing cognitive and control limitations in relation to such constructions (Burns and Deville, 2003; Machado, 1990, 1998). The paper emphasizes the importance of investigating and theorizing the particular ways in which institutional as well as individual factors increase or decrease the potential risks and the incidence of accidents.

Introduction

Actor-system dynamics (ASD) emerged in the 1970s out of early social systems analysis (Baumgartner, Burns and DeVile, 1986; Buckley, 1967; Burns, 2006a, 2006b; Burns, Baumgartner and DeVile, 1985; Burns and others, 2002). Social relations, groups, organizations, and societies were conceptualized as sets of inter-related parts with internal structures and processes. A key feature of the theory was its consideration of social systems as open to, and interacting with, their social and physical environments. Through interaction with their environment — as well as through internal processes — such systems acquire new properties and are transformed, resulting in emergent properties and evolutionary developments. Another major characteristic of the theory has entailed a conceptualization of human agents as creative (as well as destructive) transformative forces. It has also been axiomatic from the outset that human agents moral agents, shaping, reshaping, and implementing normative

1 This article — to appear in two parts — draws on an earlier paper of the authors presented at the Workshop on “Risk Management”, jointly sponsored by the European Science Foundation (Standing Committee for the Humanities) and the Italian Institute for Philosophical Studies, Naples, Italy, October 5-7, 2000. It was also presented at the European University Institute, Florence, Spring, 2003. We are grateful to Joe Berger, Mary Douglas, Mark Jacobs, Giandomenico Majone, Rui Pena Pires, and Claudio Radaelli and participants in the meetings in Naples and Florence for their comments and suggestions.

and other moral rules. They have intentionality, they are self-reflective and consciously self-organizing beings. They may choose, however, to deviate, oppose, or act in innovative and even perverse ways relative to the norms, values, and social structures of the particular social systems within which they act and interact.

Human agents, as cultural beings, are constituted and constrained by social rules and complexes of such rules (Burns and Flam, 1987). These provide the major basis on which people organize and regulate their interactions, interpret and predict their activities, and develop and articulate accounts and critical discourses of their affairs. Social rule systems are key constraining and enabling conditions for, as well as the products of, social interaction (the duality principle).

The construction of ASD has entailed a number of key innovations: (1) the conceptualization of human agents as creative (also destructive), self-reflective, and self-transforming beings; (2) cultural and institutional formations constituting the major environment of human behavior, an environment in part internalized in social groups and organizations in the form of shared rules and systems of rules; (3) interaction processes and games as embedded in cultural and institutional systems which constrain, facilitate, and, in general, influence action and interaction of human agents; (4) a conceptualization of human consciousness in terms of self-representation and self-reflectivity on collective and individual levels; (5) social systems as open to, and interacting with, their environment; through interaction with their environment and through internal processes, such systems acquire new properties, and are transformed, resulting in their evolution and development; (6) social systems as configurations of tensions and dissonance because of contradictions in institutional arrangements and cultural formations and related struggles among groups; and (7) the evolution of rule systems as a function of (a) human agency realized through interactions and games (b) and selective mechanisms which are, in part, constructed by social agents in forming and reforming institutions and also, in part, a function of physical and ecological environments.

General framework

This section identifies a minimum set of concepts essential to description and model-building in social system analysis (see figure 1 below; the following roman numerals are indicated in figure 1).

- (I) The diverse constraints and facilitators of the actions and interactions of human agents, in particular: (IA) Social structures (institutions and cultural formations based on socially shared rule systems) which structure and regulate agents and their interactions, determining constraints as well as facilitating opportunities for initiative and transformation. (IB) Physical structures which constrain as well as sustain human activities, providing, for instance, resources necessary for life and material development. Included here are physical and ecological factors (waters, land, forests, deserts, minerals, other resources). (IA, IB) Socio-technical systems combine material and social structural elements. (1A-S) and (1B-S) in figure 1 are, respectively, key social and material (or “natural”) structuring and

selection mechanisms that operate to constrain and facilitate agents' activities and their consequences; these mechanisms also allocate resources, in some cases generating sufficient "payoffs" (quantity, quality, diversity) to reproduce or sustain social agents and their structures; in other cases not.

- (II) Population(s) of interacting social agents, occupying positions and playing different roles *vis-a-vis* one another in the context of their socio-structural, socio-technical, and material systems. Individual and collective agents are constituted and regulated through such social structures as institutions; at the same time, they are not simply robots performing programs or implementing rules but are adapting, filling in particulars, and innovating.
- (III) Social action and interaction (or game) processes that are structured and regulated through established material and social conditions.² Social actors (individuals and collectives together with interaction processes make up human agency.
- (IV) Interactions result in multiple consequences and developments, intended and unintended: productions, goods, wastes, and damages as well as impacts on the very social and material structures that constrain and facilitate action and interaction. That is, the actions IVA and IVB operate on the structures IA and IB, respectively. Through their interactions, social agents reproduce, elaborate, and transform social structures (for instance, institutional arrangements and cultural formations based on rule systems) as well as material and ecological conditions.

In general, while human agents — individuals as well as organized groups, organizations and nations — are subject to institutional and cultural as well as material constraints on their actions and interactions, they are at the same time active, possibly radically creative/destructive forces, shaping and reshaping cultural formations and institutions as well as their material circumstances. In the process of strategic structuring, agents interact, struggle, form alliances, exercise power, negotiate, and cooperate within the constraints and opportunities of existing structures. They change, intentionally and unintentionally — even through mistakes and performance failures — the conditions of their own activities and transactions, namely the physical and social systems structuring and influencing their interactions. The results entail institutional, cultural, and material developments but not always as the agents have decided or intended.

This model conceptualizes three different types of causal drivers, that is factors that have the capacity to bring about or neutralize or block change (that is, to change or maintain conditions or states of the social as well as natural worlds). This multi-causal approach consists of causal configurations or powers that affect the processes and outcomes of human activities and developments (Burns and Dietz, 1992a). Three causal forces are of particular importance and make up the "iron triangle" of human agency, social structure, and environment. In particular:

2 Action is also constrained and facilitated by the responses of others who have the power to positively or negatively sanction, to persuade or inform. In other words, the agency of some actors affects the ability of other actors to exercise their own agency. In the extreme, powerful actors can severely restrict the agency of others in selected domains of social life.

- (1) human agency causal matrix. Actors operate purposively to effect their conditions; through their actions, they also have unanticipated and un-intended impacts. As indicated in the diagram, actions and outcomes are diverse (see III-IV in figure 1). Actors direct and influence one another; for instance through affecting one another's cognitive and normative orientations. Agential causality can operate either on process levels (that is, within an institutional frame) as when those in positions of authority and power can influence others or make particular collective decisions within given norms and other constraints (see III in figure 1).³
- (2) Social structures (norms, values, and institutions) also generate a type of causal force (IA-S). They pattern and regulate social actions and interactions and their consequences; however, ASD theory recognizes, as stressed earlier, that human agents may, under some conditions, ignore or redirect these arrangements, thereby neutralizing or transforming the causal forces of institutions and cultural formations.
Our emphasis here is on "internal" agents and social structures. Of course, "external" agents and institutions typically impact on activities and developments within any given social system. But these are special cases of factors (1) and (2) referred to above.
- (3) The natural and ecological causal complex is the third type of causal force (IB-S). Purely environmental or "natural" forces operate "selecting" and structuring (constraining/facilitating) human actions and interactions — at the same time that human agents have to a greater or lesser extent impacts, in some cases massive impacts, on the physical environments on which humanity and other species depend for survival, as suggested in the model.

Technology and socio-technical systems in the ASD framework

Technology, as a particular type of human construction, is defined in ASD as a complex of physical artifacts along with the social rules employed by social actors to understand, utilize and manage the artifacts. Thus, technology has both material and cultural-institutional aspects. Some of the rules considered are the "instruction set" for the technology, the rules that guide its effective operation and management. These rules have a "hands on", immediate practical character and can be distinguished from other rule systems such as the culture and institutional arrangements of the socio-technical system in which the technology is imbedded. The socio-technical system encompasses laws and normative principles as well as other rules, specifying

3 Agency can function also on structural levels, operating upon institutional frameworks, socio-technical systems, and societal arrangements, that is, the exercise of meta-power (IV-A and IV-B). The exercise of meta-power involves, among other things, establishing incentives structures and opportunity and constraining structures for agents who have or potentially have dealings with one another (Burns, Baumgartner and DeVille, 1985). Meta-power actors are structuring, allocating, selecting in ways that maintain (or reproduce) and change social structures but also impact to a greater or lesser extent on the physical environment and ecosystems.

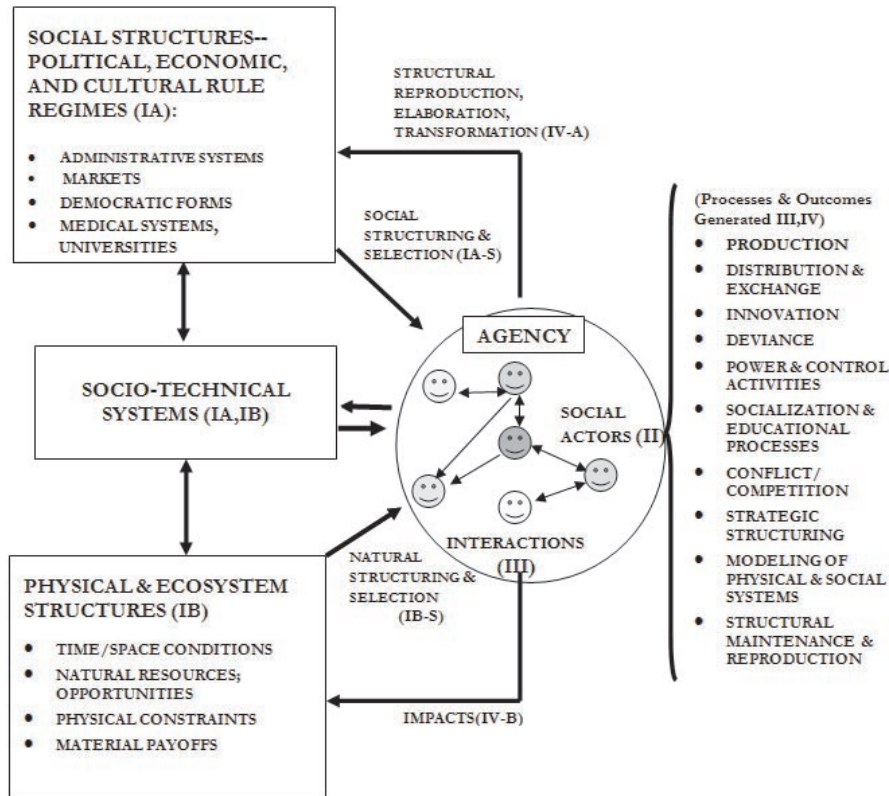


Figure 1 General ASD model: the structuring powers and socio-cultural and material embeddedness of interacting human agents

the legitimate or acceptable uses of the technology, the appropriate or legitimate owners and operators, the places and times of its use, the ways the gains and burdens (and risks) of applying the technology should be distributed, and so on. The distinction between the specific instruction set and the rules of the broader socio-technical system are not rigid, but the distinction is useful for many analytical purposes. A socio-technical system includes then the social organization (and, more generally, institutional arrangements) of those who manage, produce, and distribute its “products” and “services” to consumers and citizens as well as those (regulators, managers, and operatives) who deal with the hazards of its use and its social, health, and environmental impacts.

Such socio-technical systems as, for example, a factory, a nuclear power plant, an air transport or electricity system, organ transplantation system (Machado, 1998), money systems (Burns and DeVille, 2003), or telecommunication network

consist of, on the one hand, complex technical and physical structures that are designed to produce, process, or transform certain things (or to enable such production) and, on the other hand, institutions, norms, and social organizing principles designed to regulate the activities of the actors who operate and manage the technology. The diverse technical and physical structures making up parts of a socio-technical system may be owned and managed by different agents. The knowledge including technical knowledge of these different structures is typically dispersed among agents in diverse professions. Thus, a variety of groups, social networks, and organizations may be involved in the design, construction, operation, and maintenance of complex socio-technical systems. The diverse agents involved in operating and managing a given socio-technical system require some degree of coordination and communication. Barriers or distortions in these linkages make for likely mal-performances or system failures. Thus, the “human factor” explaining mis-performance or breakdown in a socio-technical system often has to do with organizational and communicative features difficult to analyze and understand (Burns and Dietz, 1992b; Burns and others, 2002; Vaughn, 1999).

Technologies are then more than bits of disembodied hardware; they function within social structures where their usefulness and effectiveness is dependent upon organizational structures, management skills, and the operation of incentive and collective knowledge systems (Baumgartner and Burns, 1984; Rosenberg, 1982: 247-8), hence, the importance of in our work of the concept of socio-technical system. The application and effective use of any technology requires a shared cognitive and judgment model or paradigm (Burns and others, 2002; Carson and others, 2009). This model includes principles specifying mechanisms that are understood to enable the technology to work and its interactions with its physical, biological, and socio-cultural environments. Included here are formal laws of science as well as many ad-hoc “rules of thumb” that are incorporated into technology design and use.

The concept of a socio-technical system implies particular institutional arrangements as well as culture. Knowledge of technology-in-operation presupposes knowledge of social organization (in particular, knowledge of the organizing principles and institutional rules — whether public authority, bureaucracy, private property, contract law, regulative regime, professional skills and competencies, etc. (Machado, 1998)). Arguably, a developed systems approach can deal with this complexity in an informed and systematic way. The model of a socio-technical system should always include a specification and modeling not only of its the technology and technical infrastructure but of its social organization and the roles and practices of its managers, operatives, and regulators and the impacts of the operating system on the larger society and the natural environment.

In the following sections, we apply ASD systems theory to the analysis of hazardous technologies and socio-technical systems with some likelihood of leading to accidents, that is, risky systems, and their more effective management and regulation.

Risky innovations and risky systems

Risky technologies and socio-technical systems are those which have the potential (a certain (even if very low) likelihood, to cause great harm on those involved, possibly partners or clients, third parties, other species, and the environment. Some risky systems have catastrophic potential in that they are capable in case of a performance or regulatory failure to kill hundreds or thousands, wiping out species, or irreversibly contaminating the atmosphere, water, or land.

There are a number of potentially hazardous systems which are designed and operated to be low risk systems, for instance air traffic control systems. When successful, they are characterized by a capacity to provide high qualities of services with a minimum likelihood of significant failures that would risk damage to life and property (LaPorte, 1978, 1984; LaPorte and Consolini, 1991). However, they are often costly to operate. The point is that humans construct many hazardous systems (see later) that have the potential to cause considerable harm to those involved, third parties, or the natural environment. The key to dealing with these risks is “risk manageability” — the extent that hazards can be managed, effectively regulated.

Some technologies and socio-technical systems are much more risky than others, e.g., systems of rapid innovation and development (Machado, 1990; Machado and Burns, 2001) entail unknown hazards or hazards whose likelihood are also unknown. This has to do not only with the particular hazards they entail or generate, but with the level of knowledge about them and the capacity as well as commitment to control the systems. A hierarchical society with a powerful elite may have a vision or model which it imposes, ignoring or downplaying key values and considerations of particular weak groups or even overall sustainability. In other words, their projects and developments generate risks for weak and marginal groups, and possibly even for the sustainability of the society itself over the long-run. Typically, this may be combined with suppression of open discussion and criticism of projects and their goals. Even a highly egalitarian society may generate major risks, for instance, when agents in the society are driven to compete in ways which dispose them to initiate projects and transformations that are risky to the physical and social environment. In this sense, particular institutional arrangements such as those of modern capitalism⁴ effectively drive competitiveness and high innovation levels (Burns, 2006a). For instance, in the chemical sectors, new products and production processes tend to be generated that without adequate regulation are risky for, among others, workers, consumers, the environment, and long-term system sustainability. Risky systems arise also from the fact that institutional arrangements and professional groups are inevitably biased in terms of the values they institutionalize and realize through their operations. They entail definitions of reality and social

4 Particular types of markets, super-powerful transnational corporations, regulative regimes, and many new technologies being introduced are complex, dynamic systems entailing a variety of different risks.

controls that may block or prevent recognizing and dealing with many major types of risks from technologies and technological developments (although “risk analysis” and “risk management” are very much at the forefront of their discourses).

Many contemporary developments are characterized by seriously limited or constrained scientific models and understandings of what is going on and what is likely to occur. At the same time many of these developments are revolutionizing human conditions, and we are increasingly witnessing new discourses about bounded rationality and risky systems. For instance, areas of the life sciences and medicine are inspired by the modernist claims to ultimate knowledge and capacity to control human conditions (Kerr and Cunningham-Burley, 2000; Machado and Burns, 2001).⁵ Consider several such developments in the area of biomedicine that have led to unintended consequences and new legal and ethical challenges of regulation. All of them have been launched with much promise but they have entailed complex ramifications and the emergence of major issues and problems not initially recognized or considered.

- (1) Life support technologies — life support entails a whole complex of technologies, techniques, and procedures organized, for instance, in intensive care units (ICUs). Initially, they were perceived as only a source of good — saving lives. Over time, however, they confronted hospitals, the medical profession, the public, and politicians with a wide variety of new problems and risks. The development has generated a variety of problematic (and largely unanticipated) conditions. The increasing power of these technologies has made death more and more into a construction, a “deed.” The cultural implications of withholding and withdrawing treatment (“passive euthanasia”), other forms of euthanasia, and increasingly “assisted suicide,” etc. have led to diverse ethical dilemmas and moral risks and are likely to have significant (but unknown for the moment) consequences for human conceptions and attitudes toward death (Machado, 2005, 2009).
- (2) The New Genetics — the new genetics (Kerr and Cunningham-Burley, 2000; Machado and Burns, 2001), as applied to human health problems, involves an alliance of the biotechnology industry, scientists and clinicians from an array of disciplinary backgrounds, and policy-makers and politicians concerned with health care improvement as well cost reductions. Genetic tests providing risk estimates to individuals are combined with expert counseling so

5 World Wide Web developments provide other examples. Internet was initially developed by academicians. Later, the idea emerged and spread of the usefulness of internet for multiple purposes. It was expected: that it would function as a pure source of unlimited information and knowledge development; that small companies and cooperatives could gain from safe access to global networks free and ideal transcultural exchange and learning could take place. But the emerging reality was somewhat different, a mixed picture. Among the unanticipated developments: internet as a zone of risk (e.g., from hackers) the spread of misleading information. For instance, someone may report an accident or political crisis. Others naively spread this initial report, making for a process with a non-rational life of its own; criminal activity such as child pornography, deceitful schemes, etc.; violent political groups, neo-nazis, terrorists, etc.

that those at health risk can plan their life choices more effectively. Also, the supply of information about and control over their or their offspring's genetic makeup is heralded as a new biomedical route not only to health improvement but to liberation from many biological constraints. However, its development and applications is likely to lead to a number of dramatic changes, many not yet knowable at this point in time:⁶ thus, there is emerging new conceptions, dilemmas, and risks relating health and illness. And there are increasing problems (and new types of problem) of confidentiality and access to information and protection of the integrity of individuals. Genetic testing offers the potential for widespread surveillance of the population's health by employers, insurance companies and the state (via health care institutions) and further medicalisation of risk (Kerr and Cunningham-Burley, 2000: 284).⁷ Finally, there are major risks of 'backdoor eugenics' and reinforcement of social biologism as a perspective on human beings (Machado, 2007).⁸

- (3) Xenotransplantation — xenotransplantation (transplantation of organs and tissues from one species, for instance pigs, to another, mankind) began to develop in the late 1980's as a possible substitute to organ replacement from human donors with the purpose of creating an unlimited supply of cells and organs for transplantation (Hammer, 2001). According to some observers, there are many uncertainties and risks not just for the patient but also for the larger community. — The risk of interspecies transmission of infectious agents via xenografts has the potential to introduce infectious agents including endogenous retroviruses into the wider human community with unusual or new agents. Given the ethical issues involved in xenotransplantation for, among others, the "donor," the animals, and the potentials of provoking animal rights movements, the risks are not negligible. The potential of provoking animal rights movements (as in England) may reinforce a hostile social climate that spill over and affect

6 Already there have occurred scandals and public issues. In September 1999, there was a scandal with the death of a 18 year old patient, Jesse Gelsinger from a reaction to gene therapy at the University of Pennsylvania Institute of Human Gene Therapy. Following this case, further inspections conducted by the FDA of gene therapy trials resulted in the closure and suspension of several of those clinical trials. The reason: in many of those trials the researchers were not reporting the serious adverse events suffered by research subjects. Less than 5% (37 of 970) of the serious adverse events in these gene therapies were reported (see Thompson, 2000). Thus, in addition to demand additional reporting requirements from research groups and to promote Gene Transfer Safety symposia — and in order to restore public confidence, the FDA has proposed that all researchers doing human trials of gene therapy and xenotransplantation will post "safety" information about the clinical trials, such as side effects, adverse reactions etc. in the FDA web page (www.fda.gov).

7 Such tests are useful to employers, insurance companies, and health authorities and welfare services, since they would allow them to limit their liabilities, target future services for specific 'at risk' groups, and emphasize personal responsibility for disease alleviation and prevention (Kerr and Cunningham-Burley, 2000: 289).

8 A new biological language is being developed. It is more politically correct than older languages such as that of "racism." Also, cognitive abilities" replaces the older and politically fraught concept of "intelligence." New, broader continuums of disease have been established (such as the "schizophrenic spectrum" (Kerr and Cunningham-Burley, 2000: 296).

other areas not just concerning animal welfare but also biotechnology and the important use of animals in bio-medical testing.⁹

- (4) Globalized industrial food production — today, an increased proportion of the fruits, vegetables, fish, and meats consumed in highly developed countries is grown and processed in less technologically developed countries. The procedures to process food (e.g., pasteurization, cooking, canning) normally ensure safe products. However, these processing procedures may fail in some less developed contexts. For instance, increased outbreaks of some infectious diseases are associated with animal herds (pigs, cattle, chickens). An important factor in these outbreaks is the increasing industrialization of animal-food production in confined spaces in many areas of the world that has propelled the creation of large-scale animal farms keeping substantial number of, for example, pigs or chickens in highly confined spaces. These conditions are commonly associated with a number of infectious outbreaks and diseases in the animal population, many of them a threat to human populations. Not surprisingly, this also explains in part the widespread use of antibiotics in order to avoid infections and to stimulate growth in these animal populations (increasing, however, the risk of antibiotic resistant infections in the animals and humans) (Editorial, 2000).

The existing nationally or regionally based food security and health care infrastructures are having increasing difficulty in effectively handling these problems. Earlier, people were infected by food and drink, locally produced and locally consumed — and less likely to spread widely.

- (5) Creation of many large-scale, complex systems — in general, we can model and understand only to a limited extent systems such as nuclear-power plants or global, industrial agriculture,¹⁰ global money and financial systems, etc. As a result, there are likely to be many unexpected (and unintended) developments. What theoretical models should be developed and applied to

9 “Techno-Utopianism”; “Who plays God in the 21st century?” See Turning Point Project. <http://www.turnpoint.org/>

10 Increased outbreaks of infectious diseases are associated with animal herds (pigs, cattle, chickens). An important factor in these outbreaks is the increasing industrialization of animal-food production in many areas of the world that has propelled the creation of large-scale animal farms keeping substantial number of pigs or chicken for example, in concentrated spaces. These conditions are commonly associated with a number of infectious outbreaks and diseases in the animal population, many of them a threat to human populations. Not surprisingly, this also explain the widespread use of antibiotics in order to avoid infections and to stimulate growth in these animal populations (increasing the risk of antibiotic resistant infections in humans) (Editorial, 2000). Today, an increased proportion of the fruits and vegetables consumed in highly developed countries is grown and processed in less technologically developed countries. The procedures to process food (e.g., pasteurization, cooking, canning) normally ensure safe products. However, these processing procedures can fail and sometimes do. One defective product may contaminate a number of individuals spread in different countries with a global food supply we encounter the risk that (see Editorial, 2000). The existing nationally or regionally based health care infrastructures are not prepared to handle these problems. Earlier, people were infected by food and drink, locally produced and locally consumed. We see here, in connection with technological developments, the differences between exogenous dangers and risks as opposed to endogenous dangers and risks.

conceptualize and analyze such systems. What restructuring, if any, should be imposed on these developments? How? By whom? Complex systems are developed, new “hazards” are produced which must be investigated, modeled, and controlled. At the same time, conceptions of risk, risk assessment, and risk deliberation evolve in democratic societies. These, in turn, feed into management and regulatory efforts to deal with (or prevent) hazards from occurring (or occurring all too frequently). One consequence of this is the development of “risk consciousness”, “public risk discourses”, and “risk management policies”. Such a situation calls forth public relations specialists, educational campaigns for the press and public, manipulation of the mass media, formation of advisory groups, ethics committees, and policy communities — that have become equally as important as research and its applications. They provide to a greater or lesser extent some sense of certainty, normative order, and risk minimization.

Bounded knowledge and the limits of the control of complex risky technologies and socio-technical systems

Complex systems. Our knowledge of socio-technical systems — including the complex systems that humans construct — is bounded.¹¹ Consequently, the ability to control such systems is imperfect. First, there is the relatively simple principle that radically new and complex technologies create new ways of manipulating the physical, biological, and social worlds and thus often produce results that can not be fully anticipated and understood effectively in advance. This is because they are quite literally beyond the experiential base of existing models that supposedly contain knowledge about such systems. This problem can be met by the progressive accumulation of scientific, engineering, managerial, and other practical knowledge. However, the body of knowledge may grow, even if this occurs, in part, as a consequence of accidents and catastrophes. Even then, there will always be limits to this knowledge development (Burns and Dietz, 1992b; Burns and others, 2001).

The larger scale and tighter integration of modern complex systems makes these systems difficult to understand and control (Perrow, 1999; Burns and Dietz, 1992b). Failures can propagate from one subsystem to another, and overall system performance deteriorates to that of the weakest subsystem. Subsystems can be added to prevent such propagation but these new subsystems add complexity, and may be the source of new unanticipated and problematic behavior of the overall system. Generally speaking, these are failures of design, and could at least in principle be solved through better engineering, including better “human engineering”. In practice, the large scale and complex linkages between system components and between the system and other domains of society make it very difficult to adequately understand these complex arrangements. The result is not only “knowledge problems” but “control problems”, because available knowledge cannot generate

11 This section draws on Burns and others (2002); also, see Burns and Dietz (1992b).

adequate scenarios and predictions of how the system will behave under various environmental changes and control interventions.

The greater the complexity of a system, the less likely it will behave as the sum of its parts. But the strongest knowledge that is used in many cases of systems design, construction and management is often derived from the natural sciences and engineering, which in turn are based on experimental work with relatively simple and isolated systems. There is a lack of broader or more integrative representation. The more complex the system, and the more complex the interactions among components, the less salient the knowledge about the particular components becomes for understanding the whole. In principle, experimentation with the whole system, or with sets of subsystems, could be used to elucidate complex behavior. In practice, however, such experiments become difficult and complex to carry out, too expensive and risky because the number of experimental conditions required increases at least as a product of the number of components. Actual experience with the performance of the system provides a quasi-experiment, but as with all quasi-experiments, the lack of adequate controls and isolation coupled with the complexity of the system makes the results difficult to interpret. Typically competing explanations cannot be dismissed. In any case, agreement on system description and interpretation lags, as the system evolves from the state it started from at the beginning of the quasi-experiment. This is one limit to the improvements that can be made in the models, that is knowledge, of these complex systems.

When a system's behavior begins to deviate from the routine, operators and managers must categorize or interpret the deviation in order to know what actions to take. This process involves higher order rules, including rules about what particular rules to use ("chunking rules"). Because the exceptions to normal circumstances are by definition unusual, it is difficult to develop much accumulated trial and error knowledge of them. As a result, higher order rules often are more uncertain than basic operating rules, and are more likely to be inaccurate guides to how the system will actually behave under irregular conditions. This is another way in which complexity hinders our ability to develop an adequate understanding and control of the system.

Technical division of labor — designers, builders and operators of the system are often different people working in very different contexts and according to different rules with different constraints. Each may be more or less misinformed about the rule systems used by the others. Designers may define a rigid set of rules for operators, thus allowing designers to work with greater certainty about system performance. But since the system model is imperfect, these rigid rules are likely to prevent operators from adjusting to the real behavior of the system. When they do make such adjustments — that are often useful in the local context — but they are deviating, of course, from the formal rule system, and, from the viewpoint of the systems designer, can be considered "malfunctioning" components. A further factor is the length of human life and of career patterns. This makes sure that the system's original designers are often not around anymore when operators have to cope with emergent problems, failures and catastrophes. System documentation is as subject to limitations as model

building and thus assures that operators will always be faced with “unknown” system characteristics.

Problems of authority and management — a hierarchy of authority creates different socio-cultural contexts for understanding the system and differing incentives to guide action. As one moves up in the hierarchy, pressure to be responsive to broader demands, especially demands that are external to the socio-technical system, become more important. The working engineer is focused on designing a functional, safe, efficient system or system component. Her supervisor in the case of a business enterprise must also be concerned not only with the work group’s productivity, but with the highest corporate officials’ preoccupation with enterprise profitability, and the owners of capital with the overall profitability of their portfolio. Because most modern complex systems are tightly linked to the economy and polity these external pressures at higher levels can overwhelm the design logic of those who are working “hands-on” in systems design, construction and operation. In some cases, this may be the result of callous intervention to meet profit or bureaucratic incentives. In other cases it may be the result of innocent “drift”. But in either situation, the result is much the same — the operating rules or rules-in-practice are at odds with the rules that were initially designed to optimize systems design, construction, and operation.

In addition to these macro-level interactions between the complex system and the other rule governed domains of society, there are meso- and micro-level processes at work. Managerial and other cohorts must juggle and othersong with one another and accommodate each other as individuals or groups. The day to day interaction inside and often outside the workplace makes internal mechanisms of auditing and criticism difficult to sustain. The “human factor” thus enters in in the form of deviance from safe practices and miscalculations, mistakes, and failures of complex systems.¹²

A less recognized, problem is that the processes of selection acting on rules and the processes of rule transmission will not necessarily favor rules that are accurate models of the interaction between technology and the physical, biological and social worlds. Perhaps in the very long run the evolutionary epistemology of Karl Popper and Donald Campbell will produce an improved match between the rule

12 In non-industrial and especially small-scale societies, most “system” development, including technological development, entail a substantial amount of trial and error innovation. Indeed, there is probably a direct correlation between the scale of a society and the degree to which system innovation and development depends on experimentation rather than on theory. The result is that the models and much of the knowledge that guide the development and use of human constructions, including technology, tend to be rather ad-hoc and empirically based, with limited invocation of theoretical generalizations. In the modern world, and probably in most large scale societies, the systems constructed, including technological systems, often are designed not on the basis of specific models developed inductively by experimentation with prototypes, but rather from application of the rules that constitute scientific, engineering, and managerial laws or other knowledge systems which contain their own meta-rules about forms of evidence, generalization, inference and so on. While this set of generalizations has allowed a vast expansion of system development, it also results in problems associated with the limits of such models and of de-contextualized knowledge in general.

system of a culture and “truth” but there is no guarantee that this will occur in the short run in any given culture. Even relatively simple models of cultural evolution demonstrate that disadvantageous traits can persist and even increase in frequency. The existing structure of a culture may make difficult the spread of some rules that, whatever their verisimilitude, are incongruous with other existing rules. Nor is this necessarily an unconscious process. Individuals with power may favor and sustain some rules over others, whatever their actual utility or veracity in relation to the concrete world.

The bounded rationality of models — we must recognize that the idea of bounded rationality applies to models as much as to people or organizations, since models are developed and transmitted by people and organizations. Human individuals and organizations use information-processing patterns that involve heuristics and biases, simplifications, rules of thumb and satisficing in searches for answers. In addition, since many contemporary systems including technologies are too complex for any single individual to understand fully, problems in model development result from the process of aggregating individual understandings into a collectively shared model. Aggregation of individual understandings and attendant models provide cross-checks and a larger pool of understanding on which to draw, and, in that way, the collective model will be preferable to individual models, which, even if not seriously flawed in other ways, will inevitably be incomplete. But problems of group dynamics and communication interfere with accurate modeling by a group. Groups always have agendas and dynamics that are to a large degree independent of the formal tasks to which they are assigned. These perspectives and agendas mean that there are more goals “around the table” than simply developing the best possible or most accurate operative model. Alternative goals can lead to decisions about the model construction that results in a specific model less accurate than would otherwise be possible. Delphi and other group process methods were developed specifically because of these group process problems in technological decision making.

In sum, problems of individual and collective understanding and decision-making lead to flawed models (Burns and Dietz, 1992b). Formal models may often be used to get past these problems, but they cannot eliminate them entirely. Here we note that models are limited even when all the biases of individual and group decision making are purged from them. A model of a complex system is typically built by linking models of simple and relatively well understood component systems. Thus, each element of the formal model is in itself a model of reality that eventually must be a translation from an individual or group understanding to a formal, explicit, possibly mathematical, understanding of that reality. For simple processes, both the understanding and the translation into a mathematical model may be reasonably accurate and complete. But not all subsystems of a complex system are well understood. This leads to a tendency to model those processes that are well understood, usually the linear and physical features of the system, and ignore or greatly simplify elements that are not well understood. In such models, “bad numbers drive out good paragraphs”. As a result, human operators are modeled as automatons and the natural environment as a passive sink for effluent heat, materials, etc. In addition, the

long history of trial and error experimentation with the isolated components of the system, particular physical components, has led to laws describing them in ways that are reasonably precise and accurate. This halo of precision and accuracy is often transferred to other elements of the system even though they are less well researched and cannot be subject to experimental isolation. And while some of the subsystems may be relatively well understood in themselves, it is rare that the links between the systems are understood. This is because such links and the resulting complexities are eliminated intentionally in the kinds of research and modeling that characterize most physical science and engineering. Again, the halo effect applies, and a technological hubris of overconfidence and limited inquiry may result. Finally, we should note that the model used to design and control the behavior of the system is in itself a part of the system. Since it cannot be isomorphic with the system, the behavior of the model must be taken into account when modeling the system, leading to an infinite regress.

The functioning and consequences of many innovations cannot be fully specified or predicted in advance. Of course, tests and trials are usually conducted. In the case of complex systems, however, these cover only a highly selective, biased sample of situations. Performance failings in diverse, in some cases largely unknown environments, will be discovered only in the context of operating in these particular environments.¹³ Not only is it not possible to effectively identify and test all impacts (and especially long-term impacts) of many new technologies, whose functioning and consequences are difficult to specify. But there are minimal opportunities to test complex interactions. Among other things, this concerns the impact of new technologies on human populations, where typically there is great variation in people's sensitivity, vulnerability, and absorption, etc.

Of course, the critical criterion for model adequacy is whether or not the model is helpful in designing and controlling the system. A model, though inevitably incomplete and inaccurate, may be sufficiently complete and accurate to be of great practical value. But we also note that there are strong tendencies for such models to be more inaccurate and incomplete in describing some aspects of the system than others — particularly in describing complex interactions of components of the system, the behavior of the humans who construct, manage, and operate the system, and the interactions of the systems with the natural and social environments. The failure to understand the internal physical linking of the system usually calls for more sophisticated research and modeling. The failure to understand human designers, builders and operators is labeled "human error" on the part of designers, builders and operators, rather than as an error in the systems model. These failings speak for a more sophisticated social science modeling of the "human factor" in relation to complex technologies and socio-technical systems, as the next section sets out to accomplish.

13 Such environments may be generated in part through the very application of the technology.

The complexity of governance systems and regulative limitations

We have suggested here the need for more integrative approaches. This is easier said than done. Modern life is characterized by specialization and the fragmentation of knowledge and institutional domains. There is a clear and present need for an overarching deliberation and strategies on the multiple spin-offs and spill-overs of many contemporary technology developments and on the identification and assessment of problems of incoherence and contradiction in these developments.

That is, problems of integration are typical of many technological issues facing us today. National governments are usually organized into ministries or departments, each responsible for a particular policy area, whether certain aspects of agriculture, environment, foreign affairs, trade and commerce, finance, etc. Each ministry has its own history, interests and viewpoints, and its own “culture” or ways of thinking and doing things. Each is open (or vulnerable) to different pressures or outside interest groups. Each is accountable to a greater or lesser extent to the others, or the government, or the public in different ways.

Policy formulation, for example in the area of bio-diversity, cuts across several branches of a government, involves forums outside of the government or even outside inter-governmental bodies. And in any given forum, a single ministry may have its government’s mandate to represent and act in its name. This might be the ministry of foreign affairs, asserting its authority in forums which may also be the domains of other ministries (e.g., Agriculture-FAO; Environment-UNEP).¹⁴ A variety of NGOs are engaged. Consider agricultural-related bio-diversity. It is perceived in different ways by the various actors involved: for instance, (i) as part of the larger ecosystem; (ii) as crops (and potential income) in the farmer’s field; (iii) as raw material for the production of new crop varieties; (iv) as food and other products for human beings; (v) as serving cultural and spiritual purposes; (vi) as a commodity to sell just as one might sell copper ore or handicrafts; (vii) or as a resource for national development. In short, many different interest groups are in fact interested in it.

Consequently, there is substantial complexity and fragmentation of policy-making concerning bio-diversity. As (Fowler, 1998: 5) stresses: “Depending on how the ‘issue’ is defined, the subject of agro-biodiversity can be debated in any of a number of international fora, or in multiple fora simultaneously. It can be the subject of debate and negotiation in several of the UN’s specialized agencies, inter alia, the Food and Agriculture Organization (FAO), the UN Development Programme (UNDP), the UN Environment Programme (UNEP), the UN Conference on Trade and Development (UNCTAD), the World Health Organization (WHO), the International Labour Organization (ILO), the UN Economic, Social and Cultural Organization (UNESCO), the World Trade Organization (WTO), the UN’s Commission on Sustainable Development, or through the mechanism of a treaty such as the

14 Yet, the foreign ministry typically lacks the technical expertise of the specialist ministries, and this is one of the grounded driving competition among ministries from the same country.

Convention of Biological Diversity.” Each might assert a logical claim to consider some aspect of the topic; thus, government agencies might pursue their interests in any of these fora, choosing the one, or the combination, which offers the greatest advantage. Some ministries within some governments — may consider it useful to try to define the issues as trade issues, others as environmental issues, and still others as agricultural or development issues. But in each case a different forum with different participants and procedures would be indicated as the ideal location for struggle, according to the particular framing (Fowler, 1998: 5).

Fowler (1998: 5) goes on to point out: “The multiplicity of interests and fora, and the existence of several debates or negotiations taking place simultaneously, can tax the resources of even the largest governments and typically lead to poorly coordinated, inconsistent and even contradictory policies. To some extent, contradictory policies may simply demonstrate the fact that different interests and views exist within a government. Contradictions and inconsistencies may, amazingly, be quite local and purposeful. But, in many cases, ragged and inconsistent policies can also be explained in simpler terms as poor planning, coordination and priority setting. More troubling is the fact that discordant views enunciated by governments in different negotiating fora can lead to lack of progress or stalemate in all fora.”

This case, as well as many others, illustrates the complexity of policymaking and regulation in technical (and environmental) areas. Further examples can be found in numerous areas: energy, information technology, bio-technologies, finance and banking, etc. The extraordinary complexity and fragmentation of the regulation environment make for high risky systems, as controls work at cross-purposes and breakdown. There is an obvious need for more holistic perspectives and long-term integrated assessments of technological developments, hazards, and risks.

Toward a socio-technical systems theory of risky systems and accidents¹⁵

Our scheme of complex causality (see figure 1), enables us to identify ways in which configurations of controls and constraints operate in and upon hazardous systems to increase or decrease the likelihood of significant failures that cause, or risk causing, damage to human life and property as well as to the environment. For example, in a complex system such as an organ transplant system, there are multiple phases running from donation decisions and extraction to transplantation into a recipient (Machado, 1998). Different phases are subject to more or less differing laws, norms, professional and technical constraints. To the extent that these controlling factors operate properly, the risks of unethical or illegal behavior or organ loss as well as transplant failure are minimized — accidents are avoided — and people’s lives are saved and their quality of life is typically improved in ways that are considered legitimate. Conversely, if legal,

15 Like Perrow (1999) and Rosa (2005), our approach emphasizes social systems, organizations, and institutions as key objects of analysis in determining risk and catastrophe. The field of risk analysis has been, and continues to be, dominated by psychological reductionism (Rosa, 2005: 230).

ethical, professional, or technical controls breakdown (because of lack of competence, lack of professional commitment, the pressures of contradictory goals or groups, organizational incoherence, social tensions and conflicts among groups involved, etc.), then the risk of failures increase and either invaluable organs are lost for saving lives or the transplant operations themselves fail. Ultimately, the entire system may be de-legitimized and public trust and levels of donation and support of transplantation programs decline.

Our ASD conceptualization of risky systems encompasses social structures, human agents (individuals and collective), technical structures, and the natural environment (in other analytic contexts, the social environment is included) and their interplay (see figure 1). Our classification scheme presented below in table 1 uses the general categories of ASD: agency, social structure, technical structure, environment, and systemic (that is the linkages among the major factor complexes).

Section 2 pointed out that the theory operates with configurations of causal factors (that is, the principle of multiple types of qualitatively different causalities or drivers applies):¹⁶ in particular, the causal factors of social structure, technical structure, and human agency (social actors and their interactions) as well as environmental factors (physical and ecosystem structures) and the interplay among these complexes (see figure 1). These causal factors are potential sources of failings, or the risks of failure, in the operation or functioning of a hazardous technology system., that is, the likelihood of accidents with damage to life and property as well as to the environment (even for low-hazardous systems, there are of course problems of failures).

Below we exhibit in table 1 how ASD theory enables the systematic identification of the values of variables which make for a low-risk or, alternatively, a high-risk operation of the same socio-technical system.

Table 1 also indicates how our models enable a bridging of the large gap between Perrow's framework and that of LaPorte on highly reliable risky systems (also, see Rosa (2005) about the challenge of this bridging effort of normal occupants of complex systems). On the one hand, Perrow has little to say about agential forces and also ignores to a great extent the role of turbulent environments. On the other hand, LaPorte does not particularly consider some of the technical and social structural features which Perrow emphasizes such as non-linear interactions and tight-coupling (Perrow, 1999). The ASD approach considers agential, social and technical structural, and environmental drivers and their interplay. It allows us to analyze and predict which internal as well as external drivers or mechanisms may increase — or alternatively decrease — the risk of accidents and how they do this, as suggested in the Table below. For the sake of simplifying the presentation, we are making only a rough, dichotomous distinction between low-risk and high-risk hazardous systems. In an elaborated multi-dimensional

16 We cannot elaborate here on this concept of causality but only to point out that a major "causal factor" identified by Leveson (2004) is that of design which has a proper place in the Pantheon of causal analysis, although it is not usually referred to, at least in the social sciences. But in our systems perspective, it is a major factor.

Table 1 The multiple factors of risk and accident in the case of hazardous systems: a socio-technical system perspective

<i>Agency (and its factors)</i>		
Recruitment	Stringent selection of capable and committed managers and operatives.	High-risk hazardous systems: operate relatively less safely according to established measures of failures and with a greater likelihood of causing or risking damage to life and property (includes also Perrow cases of "normal accidents")
Training, socialization	Careful, stringent.	
Sustaining management and operatives skill & commitment	Continual training and attention to commitment of managers and operatives.	Casual.
Professionalism, and other collective disciplinary factors	Intensive, high professional commitment.	Little or no attention to continuing education or levels of commitment committed.
Organizational and management commitment to safety values and derived regulatory measures and their implementation(a "culture of safety")	Strong commitment to safety values, giving them priority over cost reductions and profitability — and consistently implemented.	Weak or non-existent professionalism.
Level of reflection on limits of "rationality", for instance, Cognitive Factors as dimensions of bounded rationality (H. A. Simon, Karl E. Weick).	High awareness from recruitment, training, regulation, and the "culture of safety".	Relatively weak or compromised commitment to safety values, e.g., managers are oriented mostly to cost reductions and profitability: operative norms and practices of the "safety culture" are weak or unreliable.
Limitations of model accuracy and reliability		Low awareness or either not understood or simply ignored.
Level of system knowledge	Generally shared recognition and knowledge about key processes.	
<i>Social structural factors</i>		Unrecognized or unknown key processes.
Interaction conditions	Relatively low complexity, linearity — or complexity successfully modeled, monitored, and regulated.	Simple systems but poorly understood or modeled. In the case of highly complex, non-linear systems, a lack of relevant knowledge, modeling, monitoring, and controls. That is, potential hazardous interactions not known or ignored.
Level of social coherence and consistency	Relatively compatible and consistent values, rule systems, role relationships	Inconsistent, contradictory values, rule systems, role relationships
Level of social integration among professional, occupational groups and also other social groupings, such as ethnic, religious, etc.	High integration, or high capacity to deal with blocked or difficult communication and cooperative links.	Low integration (cleavages, tensions, and conflicts) which block or make difficult communication and cooperative links; low capacity to manage these.
Regulatory processes	Informed, coherent regulation.	Uninformed and incoherent regulation.
2nd order regulatory processes	Updating and revision of regulatory data, models, and practices.	Little or no updating and revision. Minimum learning.
<i>Technical structural factors</i>		
System components	Highly reliable components involved in critical system performance.	Low reliable components.
Technical linkages and interaction	Components are properly linked, maintained, and adjusted or adapted as required.	Component linkages are missing, faulty, unreliable, or harmful.
<i>Environmental factors</i>		
Degree of stability or turbulence	Highly stable context or, if unstable, the performance system is effectively buffered or adapted to environmental conditions.	Turbulent context without sufficient capacity to buffer or adapt (consequently, disruption of the socio-technical system).
<i>Systemic factors (1)</i>		
Level of system integration (the degree to which parts interact and fit together)	High integration. Information, communication and material flow linkages are maintained and operate appropriately.	Low integration. One or more linkages of information, communication, or material flow is broken or functions inappropriately. (2)
Agency-structure compatibility	Agents and their interactions are compatible with the social structural specifications (such as role and role relationships) and controls.	Agents and their interactions are incompatible with the social structural specifications and controls.
System-environment relations	Either Simple, orderly or buffered relations.	Turbulent relations, with which management and operatives are unable to predict or adapt and risk failures and accidents.

(1) Leveson (2004) stresses the importance of system design and the various ways in which key controls, linkages, and component properties may be missing, or incorrectly functioning.

(2) Even a previously integrated system may undergo change where one component or sub-system changes without attention or consideration of the effects on other components or subsystems, and linkages are broken or distorted in their performance.

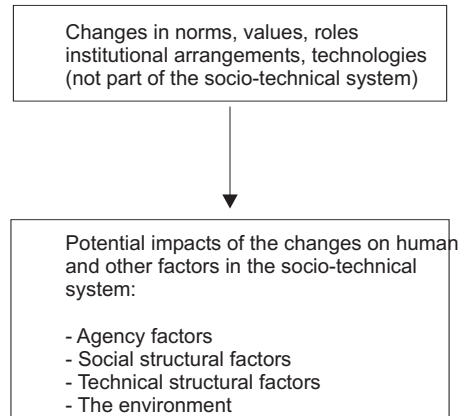


Figure 2 Impact of societal changes on the "human" and other factors of socio-technical systems

space, we would show how the scheme can be used to analyze and control each factor in terms of degree of risk it entails to the functioning and potential failure of the functioning of a hazardous socio-technical system, which can cause harm to life, property, and the environment.

Performance failures in risky systems must be comprehended in social as well as physical systems terms encompassing the multiple components and the links among components of subsystems including individual and collective agents, social structures such as institutions and organizations, and material structures including the physical environment. Risky or hazardous behavior results from the inadequate or missing controls or constraints in the system (Leveson, 2004).

A basic principle of ASD systems analysis is that effective control of major performance factors (processes as well as conditions) will reduce the likelihood of failures and accidents and make for low-risk system performance, even in the case of highly hazardous systems (that is, LaPorte's type of model). But the absence or failure of one or more of these controls and constraints will increase the risk of failure or accident in the operating system (notice that Perrow's cases are included here in a much broader range of cases, as suggested in table 1 above, column 3). As LaPorte and colleagues (LaPorte, 1984; LaPorte and Consolini, 1991) emphasize on the basis of their empirical studies of highly reliable systems, design, training, development of professionalism, redundant controls, multi-level regulation and a number of other organizational and social psychological factors can effectively reduce the risk levels of hazardous systems.

Accidents in hazardous systems occur whenever one or more components (or subsystems) of the complex socio-technical system fail, or essential linkages breakdown or operate perversely. Such internal as well as external "disturbances" fail to be controlled or buffered (or adapted to) adequately by the appropriate controllers.

The table provides a more or less static picture, but more dynamic considerations and analyses readily follow from the ASD model (see figure 2).

Thus, one can model and analyze the impact of social, cultural, and political factors on the institutional arrangements and the behavior of operatives, managers, and regulators of hazardous socio-technical systems. The following points are a selection of only a few of the possible social systemic analyses implied.

- (I) Levels of knowledge and skills of managers, operatives and regulators may decline with respect to, for instance, the socio-technical system and its environmental conditions and mechanisms. Innovations may be introduced in the socio-technical system with effects which go far beyond established knowledge of operatives, management, and regulators.
- (II) Levels of professionalism and commitment may decline because of problems (including costs) of recruitment, training, or further education of managers and operatives. Or, the decline may take place because of changes in values and norms in society, for instance, there emerges an increased emphasis on economic achievements, cost-cutting and profitability, increasing the likelihood of risk-taking and accidents. That is, values and goals unrelated to minimizing the risk of accidents are prioritized over safety.
- (III) There occurs declining capability or interest in understanding, modeling, or managing the "human factor".
- (IV) Highly competitive systems, such as capitalism, drive innovation beyond the full knowledge and skills of operatives, managers, and regulators.
- (V) Local adaptations or innovations are not communicated to others outside the local context, so that the stage is set for incompatibilities, misunderstandings, and accidents. The same type of contradictory development may occur between levels as well, that is, for instance, changes at the top management level are not communicated to or recognized by subordinate groups. Or, vice versa, shifts occur in operative groups' beliefs and practices that impact on risk-taking and safety without this being communicated to or perceived by managers and/or regulators.
- (VI) As stressed earlier, many interactions (or potential interactions) in a system or between the system and its environment are not recognized or adequately modeled. Indeed, some could not have been modeled or anticipated. They are emergent (Burns and DeVille, 2003; Machado and Burns, 2001).¹⁷ Thus, new types of hazards, vulnerabilities, and human error emerge, which the established paradigm of professionalism, training, and regulation fails to fully understand or take into account, for instance, the complex relationships resulting from increased use of automation combined with

17 This fundamental aspect has been investigated by Burns and Deville (2003) with respect to money and financial systems and by Machado in high-tech medicine (in particular, the New Genetics in medicine (Machado and Burns, 2001) and explains why regulatory instruments and regulatory institutions ultimately fail with respect to the dynamics and innovativeness of human agents and their many splendid creations.

human involvement make for new types of hazards, vulnerability, and human error (Leveson, 2004: 2-3).

- (VII) The speed of innovation — and the diversity of innovations — means that there is less time to test and learn about the various “frankensteins” and even less to find out about their potential interactions and risks (Marais, Dulac and Leveson, 2004).

In sum, the ASD system conceptualization encompasses social structures, human agents (individuals and collective), and physical artifacts as well as the environment, especially the natural environment as well as their interplay (in other analytic contexts, the social environment is included). Understanding and reducing the risks associated with hazardous systems entails (see Leveson, 2004; Leveson and others, 2009; and Marais, Dulac and Leveson, 2004:

- identifying the multiple factors, including the many “human factors” that are vulnerable or prone to breakdown or failure, resulting in accidents that cause harm to life and property as well as to the environment;
- establishing and maintaining appropriate controls and constraints — among other things, recognizing and taking measures to deal with breakdowns in controls (through human and/or machine failings);
- dealing with changes in and outside of the system which may increase hazards or introduce new ones as well as increase the risks of accidents (this would include monitoring such developments and making proactive preparations) (see figure 2).

The ASD approach advocates the development of empirically grounded and relevant theorizing, drawing on the powerful theoretical traditions of social systems theory, institutionalism and cognitive sociology, for instance in investigating the cognitive and normative bases of risk judgment and management. It provides a social science theory when enables in a systematic way the understanding, analysis, and control (that is, risk management) of complex hazardous socio-technical systems in order to prevent or minimize accidents, in particular, the role of “human factors”.

A social systems perspective re-orientes us from away from reductionist approaches toward more systemic perspectives on risk: social structure including the institutional arrangements, the cultural formations, complex socio-technical systems, etc.

Several key dimensions of risky system have been specified in this article, for instance: (i) powerful systems may have high capacities to cause harm to the physical and social environments (concerning, for instance, social order, welfare, health, etc.); (ii) hierarchical systems where elites and their advisors adhere with great confidence to, and implement, abstract models. These models have, on the one

hand, ideological or moral meanings but entail, on the other, radical impacts on the social and physical environment. Developments are decided by an elite certain of their values and truth knowledge, ignoring or neglecting other value orientations and interests, for instance, those of the populations dominated by the elite; (iii) agents may be operating in a highly competitive context which drives them to initiate projects and carry through social transformations which generate major risks for the social and physical environments. In general, “competitive systems” drive experimentation, innovation, and transformation (Burns and Dietz, 2001). That is, institutional arrangements and related social processes generate — or at least tolerate the application of — dangerous technologies and socio-technical constructions. The forms of modern capitalism combining hierarchical arrangements (the corporate structure of the firm) with institutional incentive systems (“the profit motive”) entail such risky structural arrangements. Economic agents are driven toward excessive exploitation of natural resources, risky disposal of hazardous wastes, and risky technological developments (exemplified by forms of advanced industrial agriculture which in Europe resulted in the production and spread of “mad-cow disease”). Capitalism and the modern state together with science and technology transform nature (untamed, unknown) into an environment of resources which can be defined in terms of prices, exchange-values and income flows. In very many areas, this is a risky business (and indeed, the “capitalist machine’s” ultimate limitation is arguably the physical environment and available natural resources as well as its capacity to absorb wastes) (Burns and others, 2002). Moreover, capitalist wealth is often mobilized to attack and de-legitimize the criticisms of, for instance, environmental movements and scientists that call for limiting risky socio-economic developments as well as to buy off governments that are under pressure to regulate; (iv) even if knowledge about the systems is high (which may not be the case), it is bounded knowledge (Simon, 1979). There will be unanticipated and unintended consequences which cause harm or threaten to harm particular groups, possibly society as a whole, and the environment; (v) all social systems including socio-technical systems, are subject to internal and external drivers stressing and restructuring the systems (not necessarily in appropriate or beneficial directions), making for an increase in existing risks or the emergence of entirely new risks; (vi) there is often a relatively low (but increasing) level of reflectivity about the bounded knowledge, unpredictability, and problems of unanticipated consequences. Feedback is low and deep social learning fails to take place, for instance, with respect to some of the negative impacts on the social and physical environments; (vii) such systems generate risks — possibly beyond the capacity of institutions to learn fast enough about and to reflect on the changes, their value or normative implications, their likely or possible implications for sustainability, and the appropriate strategies for dealing with them.

Modern institutional arrangements and their practices combine different types of selective processes, which may encourage (or at least fail to constrain) risky behavior, having extensive and powerful impacts on the physical and social environments, for instance, many industrial developments, transport systems, nuclear power facilities, electrical networks, airports, etc.). For instance, modern capitalism —

with its super-powers and mobilizing capabilities — generates a wide spectrum of risks. There are not only risky technologies, which in an ever-increasingly complex world cannot be readily assessed or controlled, but there are risky socio-technical systems of production, generating risky technologies, production processes, products, and practices (Machado, 1990; Perrow, 1994, 1999, 2004). In highly complex and dynamic systems, some hazards cannot be readily identified, and probabilities of particular outcomes cannot be determined and calculated. In other words, some risks cannot be known and measured beforehand. This is in contrast to cases of well-known, relatively closed technical systems, where one can determine all or most potential outcomes and their probabilities under differing conditions and, therefore, calculate and assess risks. Also, some negative consequences arise because values of concern to groups and communities are not taken into account — as a result of the prevailing power structure — or are never anticipated because of the limitations of the models utilized by power-wielders.

In conclusion, modern societies have developed and continue to develop revolutionary economic and technological powers — driven to a great extent by dynamic capitalism — at the same time that they have bounded knowledge of these powers and their consequences. Unintended consequences abound: social as well as ecological systems are disturbed, stressed, and transformed. But social agents and movements form and react to these conditions, developing new strategies and critical models and providing fresh challenges and opportunities for institutional innovation and transformation. Consequently, modern capitalist societies — characterized by their core arrangements as well as the many and diverse opponents to some or many aspects of capitalist development — are involved not only in a global struggle but a largely uncontrolled experiment (or, more precisely, a multitude of experiments). The capacity to monitor and to assess such experimentation remains severely limited, even more so with the rapid changes in global capitalism. Because the capacity to constrain and regulate global capitalism is currently highly limited, the riskiness of the system is greatly increased, raising new policy and regulatory issues. How is the powerful class of global capitalists to be made responsible and accountable for their actions? What political or regulatory forms and procedures might link the new politics suggested above to the global capitalist economy and, thereby, increase the likelihood of effective governance and regulation.

In the continuing human construction of different technologies and socio-technical systems, there are many hazards and risks, some of them indeed very dangerous and compromising of life on earth such as nuclear weapons, fossil fuel systems bringing about radical climate change, etc. Humanity only knows a small part of the mess they have been creating, and continue to create. A great many of our complex constructions are not fully predictable or understandable — and not completely controllable. An important part of the “risk society,” as we have interpreted it (Burns and Machado, 2009), are the higher levels of risk consciousness, risk discourse, risk demands and risk management. These have contributed — and hopefully will continue to contribute in the future — to more effective control of risks, or eliminating some types of risks altogether. This article has aimed to

conceptualize and enhance the regulation of the multiple “dimensions” and mechanisms of the “human factor” in risky systems and cases of accidents.

- Baumgartner, T., and T. R. Burns (1984), *Transitions to Alternative Energy Systems. Entrepreneurs, Strategies, and Social Change*, Boulder, Colorado, Westview Press.
- Baumgartner, T., T. R. Burns, and P. DeVille (1986), *The Shaping of Socio-Economic Systems*, London, England, Gordon and Breach.
- Beck, U. (1992), *Risk Society. Towards a New Modernity*, London, Sage Publications.
- Buckley, W. (1967), *Sociology and Modern Systems Theory*, Englewood Cliffs, NJ, Prentice-Hall.
- Burns, T. R. (2006a), “Dynamic systems theory”, in Clifton D. Bryant and D. L. Peck (eds.), *The Handbook of 21st Century Sociology*, Thousand Oaks, California, Sage Publications.
- Burns, T. R. (2006b), “The sociology of complex systems: an overview of actor-systems-dynamics”, *World Futures. The Journal of General Evolution*, 62, pp. 411-460.
- Burns, T. R., T. Baumgartner, and P. DeVille (1985), *Man, Decision and Society*, London, Gordon and Breach.
- Burns, T. R., T. Baumgartner, and P. DeVille (2002), “Actor-system dynamics theory and its application to the analysis of modern capitalism”, *Canadian Journal of Sociology*, 27 (2), pp. 210-243.
- Burns, T. R., T. Baumgartner, T. Dietz, and N. Machado (2002) “The theory of actor-system dynamics: human agency, rule systems, and cultural evolution”, in *Encyclopedia of Life Support Systems*, Paris, UNESCO.
- Burns, T. R., and M. Carson (2002), “Actors, paradigms, and institutional dynamics”, in R. Hollingsworth, K. H. Muller, E. J. Hollingsworth (eds.), *Advancing Socio-Economics. An Institutional Perspective*, Oxford, Rowman and Littlefield.
- Burns, T. R., and P. DeVille (2003), “The three faces of the coin: a socio-economic approach to the institution of money”, *European Journal of Economic and Social System*, 16 (2), pp. 149-195.
- Burns, T. R., and T. Dietz (1992a) “Cultural evolution: social rule systems, selection, and human agency”, *International Sociology*, 7, pp. 259-283.
- Burns, T. R., and T. Dietz (1992b), “Technology, sociotechnical systems, technological development: an evolutionary perspective”, in M. Dierkes and U. Hoffman (eds.), *New Technology at the Outset. Social Forces in the Shaping of Technological Innovations*, Frankfurt/Main, Campus.
- Burns, T. R., and T. Dietz (2001) “Revolution: an evolutionary perspective”, *International Sociology*, 16, pp. 531-555.
- Burns, T. R., and H. Flam (1987), *The Shaping of Social Organization. Social Rule System Theory and its Applications*, London, Sage Publications.
- Burns, T. R., and N. Machado (2009) “Technology, complexity, and risk: part ii: social systems perspective on socio-technical systems and their hazards”, *Sociologia, Problemas e Práticas*, Forthcoming.

- Carson, M., T. R. Burns, and D. Calvo (eds.) (2009), *Public Policy Paradigms. Theory and Practice of Paradigms Shifts in the EU*, Frankfurt/Berlin/Oxford, Peter Lang, in process.
- Editorial (2000), "Emerging infections: another warning", *New England Journal of Medicine*, 342 (17), p. 1280.
- Fowler, C. (1998), "Background and current and outstanding issues of access to genetic resources", *ESDAR Synthesis Report*.
- Hammer, C. (2001), "Xenotransplantation: perspectives and limits", *Blood Purification*, 19, pp. 322-328.
- Kerr, A., and S. Cunningham-Burley (2000), "On ambivalence and risk: reflexive modernity and the new human genetics", *Sociology*, 34, pp. 283-304.
- LaPorte, T. R. (1978), "Nuclear wastes: increasing scale and sociopolitical impacts", *Science*, 191, pp. 22-29.
- LaPorte, T. R. (1984), "Technology as social organization", Working Paper, 84 (1), (IGS Studies in Public Organization), Berkeley, California, Institute of Government Studies.
- LaPorte, T. R., and P. M. Consolini (1991), "Working in practice but not in theory: theoretical challenges of 'high reliability organizations'", *Journal of Public Administration Research and Theory*, 1, pp. 19-47.
- Leveson, N. (2004), "A new accident model for engineering safer systems", *Safety Science*, 42 (4), pp. 237-270.
- Leveson, N., N. Dulac, K. Marais, and J. Carroll (2009), "Moving beyond normal accidents and high reliability organizations: a systems approach to safety in complex systems", *Organization Studies*, 30 (2-3), pp. 227-249.
- Machado, N. (1990), "Risk and risk assessments in organizations: the case of an organ transplantation system", Presented at the XII World Congress of Sociology, July 1990, Madrid, Spain.
- Machado, N. (1998), *Using the Bodies of the Dead. Legal, Ethical and Organizational Dimensions of Organ Transplantation*, Ashgate Publishers, Aldershot, England.
- Machado, N. (2005), "Discretionary death: cognitive and normative problems resulting from advances in life-support technologies", *Death Studies*, 29 (9), pp. 791-809.
- Machado, N. (2007), "Race and medicine", in T. R. Burns, N. Machado, Z. Hellgren, and G. Brodin (eds.), *Makt, kultur och kontroll över invandrades livsvillkor*, Uppsala, Uppsala University Press.
- Machado, N. (2009), "Discretionary death", in C. Bryant and D. L. Peck (eds.) *Encyclopedia of Death and the Human Experience*, London/Beverly Hills, Sage Publications.
- Machado, N., and T. R. Burns (2001), "The new genetics: a social science and humanities research agenda", *Canadian Journal of Sociology*, 25 (4), pp. 495-506.
- Marais, K., N. Dulac, and N. Leveson (2004), "Beyond normal accidents and high reliability organizations: the need for an alternative approach to safety in complex systems", *MIT Report*, Cambridge, Mass.
- Perrow, C. (2004), "A personal note on normal accidents", *Organization & Environment*, 17 (1), pp. 9-14.
- Perrow, C. (1999), *Normal Accidents. Living with High-Risk Technologies*, 2nd ed., Princeton, N. J.: Princeton University Press; New York: Basic Books (Originally published in 1984 by Basic Books, New York).

- Perrow, C. (1994), "The limits of safety: the enhancement of a theory of accidents", *Journal of Contingencies and Crisis Management*, 2 (4), pp. 212-220.
- Rosa, E. A. (2005), "Celebrating a citation classic: and more", *Organization & Environment*, 18 (2), pp. 229-234
- Rosenberg, N. (1982), *Inside the Black Box. Technology and Economics*, Cambridge, Cambridge University Press.
- Simon, H. A. (1979), *Models of Thought*, New Haven, Yale University Press.
- Thompson, L. (2000), "Human gene therapy harsh lessons, high hopes", *FDA Consumer Magazine*, September-October, pp. 19-24.
- Vaughan, D. (1999), "The dark side of organizations: mistake, misconduct, and disaster", *Annual Review of Sociology*, 25, pp. 271-305.

Tom R. Burns. Lisbon University Institute, Lisbon, Portugal. Woods Institute for the Environment, Stanford University, Stanford, Calif. Uppsala Theory Circle, Department of Sociology, University of Uppsala, Box 821, 75108 Uppsala, Sweden. E-mail: tomburns@stanford.edu

Nora Machado. CIES, Lisbon University Institute, Lisbon, Portugal. Science, Technology, and Society Program, Stanford University, Stanford, Calif. Department of Sociology, University of Gothenburg, Gothenburg, Sweden (on leave-of-absence). E-mail: noramachado@gmail.com

/ / /

Tecnologia, complexidade, e risco: análise social sistémica de sistemas sociotécnicos de risco e da probabilidade de acidentes

Este artigo conceptualiza a multidimensionalidade do "factor humano" em sistemas tecnológicos de risco e em casos de acidentes. A teoria dos sistemas sociais é aplicável à análise de tecnologia perigosa e de sistemas sociotécnicos, às respectivas dinâmicas complexas e dimensões de risco. O "factor humano" é muitas vezes vagamente identificado como factor de risco em sistemas sociotécnicos de risco, particularmente quando ocorrem acidentes. Mas é usualmente visto mais ou menos como a "caixa-negra", subespecificado e subanalisado. Três objectivos fundamentais do artigo são: (1) identificar e teorizar de forma sistemática o carácter multidimensional do "factor humano" em sistemas de risco e em acidentes; (2) permitir a aplicação sistemática de conhecimento substancial da ciência social à regulação de sistemas perigosos, seus gestores e operadores bem como reguladores, especialmente relacionados com o "factor humano"; (3) servir como ferramenta de orientação para investigadores e reguladores na compilação e organização de dados sobre humanos e outros factores em sistemas de risco e acidentes. Em suma, o artigo propõe uma abordagem sistemática para analisar muitos dos diversos factores humanos de risco associados a tecnologias complexas e a

sistemas sociotécnicos, contribuindo, assim, para o conhecimento preventivo — ou para minimizar a probabilidade — de acidentes ou catástrofes.

Palavras-chave dinâmica actor-sistema, tecnologia, sistema sóciotécnico, complexidade, risco, sistema de risco, acidente, regulação e controle.

Technology, complexity, and risk: social systems analysis of risky socio-technical systems and the likelihood of accidents

This article conceptualizes the multi-dimensional "human factor" in risky technology systems and cases of accidents. A social systems theory is applied to the analysis of hazardous technology and socio-technical systems, their complex dynamics, and risky dimensions. The "human factor" is often vaguely identified as a risk factor in hazardous socio-technical systems, particularly when accidents occur. But it is usually viewed more or less as a "black box", under-specified and under-analyzed. Three key aims of the article are: (1) to identify and theorize in a systematic way the multi-dimensional character of the "human factor" in risky systems and accidents; (2) to enable the systematic application of a substantial social science knowledge to the regulation of hazardous systems, their managers and operatives as well as regulators, especially relating to the "human factor;" (3) to serve as an guiding tool for researchers and regulators in the collection and organization of data on human and other factors in risky systems and accidents. In sum, the article proposes a systematic approach to analyzing many of the diverse human risk factors associated with complex technologies and socio-technical systems, thus contributing knowledge toward preventing -- or minimizing the likelihood of -- accidents or catastrophes.

Key-words actor-system dynamics, technology, socio-technical system, complexity, risk, risky system, accident, regulation and control.

Technologie, complexité et risque: analyse sociale systémique des systèmes socio-techniques de risque et de la probabilité d'accidents

Le présent article conceptualise la multidimensionnalité du "facteur humain" dans les systèmes technologiques de risque et en cas d'accidents. La théorie des systèmes sociaux est applicable à l'analyse de la technologie dangereuse et des systèmes socio-techniques, à leurs dynamiques complexes à leurs dimensions de risque. Le "facteur humain" est souvent vaguement identifié comme un facteur de risque dans des systèmes socio-techniques de risque, notamment en cas d'accidents. Mais il est habituellement plus ou moins considéré comme la "boîte noire", sous-spécifié et sous-analysé. Cet article a trois objectifs fondamentaux : (1) identifier et théoriser systématiquement le caractère multidimensionnel du "facteur humain" dans les systèmes de risque et les accidents ; (2) permettre

l'application systématique de la connaissance substantielle de la science sociale à la régulation des systèmes dangereux, à leurs gestionnaires, leurs opérateurs et aussi leurs régulateurs, ayant un lien particulier avec le "facteur humain"; (3) servir comme outil d'orientation aux chercheurs et aux régulateurs dans la compilation et l'organisation des données sur les humains et autres facteurs dans des systèmes de risque et les accidents. En somme, l'article propose une approche systématique pour analyser la plupart des facteurs humains de risque associés aux technologies complexes et aux systèmes socio-techniques, contribuant ainsi à la connaissance préventive des accidents ou des catastrophes, ou à en minimiser la probabilité.

Mots-clé dynamique acteur-système, technologie, système socio-technique, complexité, risque, système de risque, accident, régulation et contrôle.

Tecnología, complejidad, y riesgo: análisis social sistémico de sistemas sociotécnicos de riesgo y de la probabilidad de accidentes

Este artículo conceptualiza la multidimensionalidad del "factor humano" en los sistemas tecnológicos de riesgo y en casos de accidentes. La teoría de los sistemas sociales es aplicable al análisis de la tecnología peligrosa y de los sistemas sociotécnicos, de las respectivas dinámicas complejas y dimensiones de riesgo. El "factor humano" es muchas veces vagamente identificado como factor de riesgo en sistemas sociotécnicos de riesgo, particularmente cuando ocurren accidentes. Sin embargo, es usualmente visto más o menos como la "caja-negra", subespecificado y subanalizado. Son tres los objetivos fundamentales de este artículo: (1) identificar y teorizar de forma sistemática el carácter multidimensional del "factor humano" en sistemas de riesgo y en accidentes; (2) permitir la aplicación sistemática del conocimiento substancial de las ciencias sociales a la regulación de sistemas peligrosos, sus gestores y operadores así como reguladores, especialmente relacionados con el "factor humano"; (3) servir como herramienta de orientación para investigadores y reguladores en la compilación y organización de datos sobre humanos y otros factores en sistemas de riesgo y accidentes. En suma, el artículo propone un abordaje sistemático para analizar muchos de los diversos factores humanos de riesgo asociados a las tecnologías complejas y a sistemas sociotécnicos, contribuyendo, así, para al conocimiento preventivo -- o para minimizar la probabilidad -- de accidentes o catástrofes.

Palabras-clave dinámica actor-sistema, tecnología, sistema sociotécnico, complejidad, riesgo, sistema de riesgo, accidente, regulación y control.

