CAHIERS D'ÉPISTÉMOLOGIE

**Critical instrumental rationality between spontaneity and reflexivity:
Spelling out Hayekian Critical Theory**

[The use of knowledge about society, Part II]

Christian Arnsperger

Cahier nº 2004-05   **317ᵉ numéro**

UQÀM

The use of knowledge about society
Part II

**Critical instrumental rationality between spontaneity and reflexivity:
Spelling Out Hayekian Critical Theory**

Christian Arnsperger
FNRS (Brussels) & Université catholique de Louvain (Belgium)
arnsperger@etes.ucl.ac.be

September 2004

We are—and ought to be—forevermore unable to draw up a complete blueprint of the fully emancipated society, in all its procedural and consequential details, as we might wish it would emerge here and now from our ongoing complex social process. However, can't we at least reflect on the *individual dispositions* and the *interaction mechanisms* which would be conducive to the uncontrolled emergence of an emancipated society? Some of us, a minority perhaps, might believe that it need no longer emerge because it already has (Fukuyama [1992]); for the rest of us, however, the political disasters of really existing communism, as well as the philosophical and scientific advancement of ideas about "out-of-control systems" (see e.g. Kelly [1994]), have brought to the fore today's most pressing epistemological urgency—and not surprisingly, it lacks a simple, linear formulation: we need to understand how to "harness" the complexity (Axelrod and Cohen [2000]) of a social system whose aggregate states emerge as the partly unintended result of the adaptive interaction of individuals and groups of individuals, each of whom is attempting to act and react on a model intended to "harness" the system's complexity for its specific interests. (You should not jump on this word immediately and deduce that this is a "special-interest politics" approach; more on this below.) Since, moreover, it is dawning on us that each idnvidual brain is itself a complex distributed network giving rise to emergent properties in the form of cognition and intentionality (see e.g. Varela, Thompson and Rosch [1991]), we have to face up to a somewhat daunting-looking task:

> How to comprehend the emergence of an *emancipated society* as the harnessing of the complexity of a social system S of complex interactors i, each of whom acts intentionally on a *critical theory* t_i of how S works, i.e., each of whom seeks to act on t_i so as to further her *emancipatory interests* in S?

Of course, just because the problem looks daunting doesn't mean we can escape it. Understannding how markets generate order with everyone pursuing potentially disordering purposes must have seemed to the eighteenth-century Scottish and British philosophers an equally daunting task, but they tackled it. In fact, the market-order problem is a *subspecies* of the more general problem just outlined, one which postulates noncomplex brains and a specific class of individual interests being pursued (Hayek [1976], Lindblom [2000]). Now, all proportions should be kept and there will be no claim made in this paper to the effect that the above "complex-system-of-complex-agents" problem can be in any sense solved—especially, as we

shall see, since what these agents will be taken to pursue will not necessarily and not even usually be the interests lent to individuals by Smith, Hume or, for that matter, Hayek. Indeed, *emancipation* is a much broader category of aims than what even Adam Smith's morally sentimental individual, or Hayek's culturally evolved one, have been thought to pursue. What this paper aims to do, rather, is more modestly to clarify as much as possible the main issues involved in searching for a solution to the above question.

My overarching claim, to be diffracted into various subclaims, will be that a concept of *critical individual rationality* can and must be used precisely for that purpose. From the point of view of argumentative method, this will imply using the concept of a Hayekian Critical Theory of society which I have developed elsewhere (Arnsperger [2004a, 2004b, 2004c]) but will brifely rehearse here, and fitting the notion of critical instrumental rationality into the Hayekian "rule of law" in such a way as to make Hayek's own assumptions about the "use of knowledge in society" come out looking much too limited and, in fact, harmfully restrictive. In that way, I hope to show that we need to move from the Hayekian question of use of knowledge *in* society to a more Frankfurt-School-oriented question of the use of knowledge *about* society. Only if we succeed in this conversion—which does not leave the Frankfurt School unscathed, either—will we have any chance of answering the crucial question outlined earlier.

Section 1 of the paper circumscribes the various ways in which emancipation can be conceptualized in a complex adaptive system and offer a simple taxonomy that will help to clarify subsequent issues. Section 2 then analyzes the role played in social emancipation by the key concept of critical instrumental rationality, and shows how it combines crucial intuitions both from Hayekian social science and from Frankfurt School criticism. What underlies this discussion is a specific approach I call Hayekian Critical Theory, and section 3 is an attempt to demonstrate this approach at work in the study of the complex dynamics of a social system in which interactions occur between individuals endowed with critical instrumental rationality. This makes it possible to define society non-holistically as a self-criticizing system. Finally, section 4 draws certain crucial conclusions from this approach concerning the role of social science in the shaping the dynamic path of the system. The model allows us to better understand some observable characteristics of scientific communities and, more broadly, of communities of knowledge-guided individuals.

1. Varieties of emancipation in complex social systems

1.1. *Why emancipation? On the intrinsic structure of Enlightenment reason*

Contemporary neurophysiology and philosophy of mind have deeply shaken the foundations of Kantian and, more broadly, "volunatristic" anthropology. After the psychoanalytical and linguistic turns, we are now in addition realizing that the very "seat" of our consciousness may lack spiritual and even material foundation. Consciousness and volition itself—that Holy of Holies of the Enlightenment—appear to us more and more as emergent phenomena of interconnected networks of neurons; nay, our very sense of identity and our very feelings of "self" may in fact rest on a false continuity which is, in actual fact, the result of repeated reenactments of brain connections triggered (so to speak) as "perturbations" or "irritations" by our unceasing, active involvement with a natural and social environment. Thus, *who I am* is probably a tentative narrative that emerges from a sequence of "microidentities" succeeding each other through breakdowns or crises of their associated "microworlds." The "I" is the immediate coping device holding together the *history of structural couplings* between microidentities and microworlds. (On all this, see Maturana and Varela [1987], Varela, Thompson and Rosch [1991], Varela [1999], and Lakoff and Johnson [1999].)

It is difficult to see how we could nowadays make sense of the *idea of autonomy* without heeding the results from cognitive science. Far from being only a property of a system that transforms inputs—"messages"—into outputs—"actions"—by processing its environment and possibly bending it to its interests (Cartesian version) or filtering it through its wired-in moral imperatives (Kantian version), autonomy is also a property of *internal coherence of an operationally closed system* which constantly recreates itself (*autopoiein*) through its structural couplings with various environments (Varela [1983]). In both cases, autonomy means essentially the capacity for not being made suddenly unrecognizable by inputs or perturbations; it means changing while not changing too much, too abruptly. These two perspectives need not systematically exclude each other, but they do highlight the difficulty, once one accepts that human subjects as systems are engulfed in a dynamic flow of mutually interpenetrating and interacting environments, of balancing the "passive" component of information reception and the "enactive" component of embodied cognition.

In both instances, the *puzzle of internal finality* pops up inevitably. Jean-Pierre Dupuy (2002a: 10 and 46) has argued that it can usually be—and frequently has been—resolved through a retreat into an *as if* methodology: if we remain content with making internal finality apparent or heuristically descriptive, we need not postulate any actual *telos* in the system under consideration—not even in the human subject. Dupuy himself certainly does not believe this is always a sound retreat, but he is totally correct in saying that a "subjectless philosophy of mind" has often tended to treat conscious desires and aspirations as virtual epiphenomena of neuronal self-organization; and he offers this striking insight: "Kantianism, yes, but without the transcendental subject (…). This formula applies wonderfully well to cybernetics. (…) A model of the mind, yes; of the subject, no. (…) In inventing a type of transcendental inquiry that did away with the subject, cybernetics was to greatly assist the deconstruction of the metaphysics of the subject" (Dupuy [2002a: 107]). Dupuy's strong theses around this crucial claim have not yet received sufficient attention in philosophical circles: his idea that cybernetics was not at all (as often believed) designed to make machines into subjects but to make subjects into machines—what he calls the project of "mechanization of the mind"—leads him to interpret contemporary philosophies of autonomy, most notably John Rawls's recasting of Kantian autonomy, as attempts to model human subjects as "machines" so as to simply assume away their ability to cause violence and evil and to make their *autopoiesis* serve horrendous goals (Dupuy [2002b]).

Now, what the purported silencing of metaphysics has also done—and to this Dupuy does not allude directly—is to sweep under the carpet something which Kant himself, most notably in his famed essay "Was ist Aufklärung?", saw as the hallmark of the Enlightenment, namely the *deep-seated human desire for emancipation*. One of Kant's difficulties—and here we do reconnect with Dupuy's own critique of cybernetic autonomy—was to accept that revolutionary violence might in some historical situations be the only way of breaking a deadlock and of collectively attaining freedom from oppressive social structures; this led him to his famous claim that while the French revolution was in principle one of the great emancipatory achievements of the *Aufklärung*, the individuals whose actions, decisions and interactions propelled the emergence of the revolutionary phenomenon were irrational, bloodthirsty brutes unworthy of emancipation. Nevertheless, the connection between individual rationality and the aspiration to (individual

and/or collective) emancipation from oppressive structures has been strong and enduring in the whole Enlightenment tradition.

This strong end enduring link between emancipation and the metaphysics of the subject is what impelled the leader of the Frankfurt School, Max Horkheimer, to make the ironic claim in the mid-1930s that "Metaphysics may well be proud of the newest attack upon it [by positivism and empiricist pragmatism]; it [metaphysics] has been identified with thought" (Horkheimer [1937a: 187]). The remark is ironic because as a materialist Marxist, Horkheimer did not himself harbor much affection for metaphysics; what he means, however, is that at least metaphysicians *think*. This implies that the pragmatists and the philosophers of common sense and common language use do not think according to Horkheimer—just as science does not think according to Heidegger. More precisely, they do not practice "thought" in the only way which Horkheimer deems truly worthy of the Enlightenment tradition, namely *thought aimed at constructing a theory of contemporary society with a view to analyzing the obstacles that stand in the way of further emancipation in that society*. Thus, in a probably needlessly hyperbolic and insulting bout to monopolize the "true" meaning of thought, Horkheimer conveys to us the fundamental idea that what is intrinsically required by the Enlightenment tradition is a specific type of knowledge which breaks away from what he calls "traditional" knowledge, namely knowledge based (inductively or deductively) on first-degree, uninterpreted observations and facts. To such traditional scientific inquiry he opposes "critical" theory. As James Bohman has put it very aptly,

According to Max Horkheimer's (…) definition, a theory is critical only if meets three criteria: it must be explanatory, practical, and normative, all at the same time. That is, it must explain what is wrong with current social reality, identify actors to change it, and provide clear norms for criticism and practical goals for the future. Horkheimer holds that these features distinguish "critical" theory from those "traditional" theories which only seek to mirror reality as it is. The critical theory of society, as Horkheimer defined it in his programmatic writings as Director of the Frankfurt School's Institute for Social Research, "has as its object human beings as producers of their own historical form of life"; its goal is "the emancipation of human beings from the circumstances that enslave them." (Bohman [1996:190]).

Thus, individual autonomy in the revised "metaphysical" sense promoted by Horkheimer as the heritage of the Enlightenment project consists in an individual's awareness that she is taking part in the production of her own historical form of life and that this form of life should be constantly subjected to critical scrutiny from the viewpoint of emancipation. In this perspective, Theodor Adorno wrote of a critical impulse: "The critical impulse is at one with resistance against the stiff

conformity of whatever is the prevailing opinion" (Adorno [1962: 555]). To clam the suspicion that such a statement might by thinly disguised elitism, one can advance Charles Lindblom's observation to the effect that we are all, more or less consciously, reflective individuals: "Everybody always probes. However badly, all of us engage in examining volitions, even while asleep and both consciously and unconsciously in waking hours. (…) We speculate from time to time about the patterns of our lives and changes we might attempt" (Lindblom [1990: 30]). Each individual is both subjected to social and historical limitations on what he can reflectively criticize, and able to criticize these limitations by second-order reflexiveness.

That this notion of *autonomy as reflection for emancipation* is quite different from the notion of autonomy as autopoiesis of an operationally closed system, but needs to be made compatible with it if we are to be both Enlightenment humanists and informed of current scientific advances, is one of the key claims of this paper. In line with what has been suggested—and in part carried out—by Dupuy for the recognition of evil and violence, we need to understand how *subjective aspirations to an emancipated society and subjectless social processes* can be brought together within social theory.

1.2. *Emancipation in complex systems: Harnessing complexity rather than making it go away*

It does not take much reflection to realize that although everything that has been said up to now is undoubtedly in line with what enlightened reason seems to demand, the very notion of emancipation has in fact remained rather abstract. By that, I emphatically do *not* mean what is usually meant in contemporary textbooks on political philosophy, namely that just talking about "freedom" or "liberty" is not enough, that these are underdetermined words which need to be specified by a philosophy of freedom (libertarian, liberal, and so on). This too is undoubtedly correct, but here I use the term "abstract" a bit differently: I mean that no affirmation of an aspiration to emancipation can hold water without a parallel reflection on (*i*) the conditions which, in the prevailing society, make that emancipation empirically feasible or only wishful thinking and (*ii*) the structural properties of the prevailing society which create the conditions identified in (*i*).

Usually, question (*i*) is harder to answer than question (*ii*) since awareness of empirical obstacles does not always, by itself, yield information about the degree to which these obstacles are due to stable or ephemeral properties of the society in which they arise. Both questions,

however, presuppose a society. Unless we are willing to be essentialistic about this concept of society—which I am not—we need to be able to say something sufficiently general but also sufficiently situated so as to convey the context in which our purported pulling together of subjective aspirations to an emancipated society and subjectless social processes is to take place. In a recent, groundbreaking book entitled *Harnessing Complexity* and perhaps the most vivid attempt to understand our social world on the basis of complexity since Herbert Simon's *Sciences of the Artificial* (Simon [1996]), Robert Axelrod and Michael Cohen claim that today's overarching concept for a social context is that of a *complex adaptive system*:

Whether or not we are aware of it, we all intervene in complex systems. We design new objects or new strategies for action. (…) Whether simple or sophisticated, such actions change the world and (…) lead to consequences that may be hard to imagine in advance. (…) The complexity of the world is real. We do not know how to make it disappear. (…) For us, "complexity" does not simply denote "many moving parts." Instead, complexity indicates that the system consists of parts which interact in ways that heavily influence the probabilities of later events. Complexity often results in features, called emergent properties, which are properties of the system that the separate parts do not have. (Axelrod and Cohen [2000: 1-2 and 15])

Recapitulating some of the pioneering research of John Holland (1995) and other at the Santa Fe Institute, the authors offer a rigorous definition of such a system—one which, with suitable and crucial modifications, we will be sticking to all through this paper:

Agents, of a variety of types, use their strategies, in patterned interaction, with each other and with artifacts. Performance measures on the resulting events drive the selection of agents and/or strategies through processes of error-prone copying and recombination, thus changing the frequencies of the types in the system. (Axelrod and Cohen [2000: 154])

What is the meaning of "emancipation" in a complex adaptive social system? To get ourselves started in thinking about this, the best thing to do is to follow Axelrod and Cohen's reasoning a bit further. Their whole book is centered on the crucial question of how the structural features of a system can be used to make for optimal interventions—a question which Holland equates essentially with that of "locating 'lever points' in complex adaptive systems" (Holland [1995: 5]). It is well known that activating such lever points will never modify the fact that the system composes its aggregate outcomes as emergent phenomena, but it may affect the structure (the types of interactions that produce the outcome), the pattern (in terms of the distribution of individual outcomes) or the value (by some aggregate criterion for judging the outcome) of that emergence. In line with this, Axelrod and Cohen seek to provide "a framework, a way of thinking through a complex setting that takes advantage of complexity *to generate new questions and new possibilities for action*" and this suggests "a device for channeling the complexity of a social

system into desirable change, just as a harness focuses the energy of a horse into the useful motion of a wagon" (Axelrod and Cohen [2000: 2], my italics). To that end they intend to

study important mechanisms that can be used to influence the amount of variety in a system so as to affect the balance between exploration and exploitation, to alter the structure of interactions within the system, and to adjust the way success is measured and amplified. (Axelrod and Cohen [2000: 3])

Now, nothing in these concepts *a priori* forbids us to attempt to establish meaningful connections with the Frankfurt School's intimations of critical theorizing and individual aspirations to emancipation. However, Axelrod and Cohen's otherwise fascinating book turn out to be, in that specific respect, a bit disappointing. Perhaps out of fear of being "ideological," but more probably out of a spontaneous attraction for business opportunities and large "clients" for their theory (which started as a report to the United States Defense Department on optimal recruiting activities to bolster internal creativity in the US Army), the two authors almost entirely focus their discussion on examples of organizational restructuring high-technology systems, and imperfect-competition management; when they do stray from that fairly narrow set of issues, to discuss AIDS prevention or the formation of social capital, still they remain mostly focused on issues that I would call narrowly connected with contemporary industrial market society. Taking that society to be a complex adaptive system, or rather a system composed of numerous complex adaptive subsystems, their book offers analytical tools to either make more money (by being smarter oneself or by organizing others' activities better) or to implement better social and economic policies—both, however, *within the overall structure* of industrial market society.

Happily, this critical judgment about one localized book can allow us to make substantial progress in our understanding of emancipation with a view to critical theory. What Axelrod and Cohen's analysis implicitly avows is that there are *at least three tiers or levels of emancipation* in a complex adaptive system. To clarify these levels, let me distinguish in a standard way between two elements: (*i*) an individual *course of action* given the system's overall structure—i.e., given (*a*) the directed graph(s) defining its interaction pattern(s) and (*b*) the motivations that generate flows between nodes of the graph(s); and (*ii*) the *overall structure of the system* itself, i.e., the features (*a*) and (*b*) just mentioned. Let S stand for the system, CA for "course of action," and OS for "overall structure." One way of connecting Axelrod and Cohen's idea of harnessing complexity with the Frankfurt School's idea of social criticism is then to distinguish three types of emancipation:

(I) *Opportunistic harnessing of social complexity:* Given the features (*a*) and (*b*) of *S*, take the courses of action which pursues your *unreflected* idea of freedom, i.e., the idea which the overall logic inherent in (*a*) and (*b*), as well as your location on the graph, spontaneously attributes to you. This category consists in each individual asking herself, "How can I be freer in the social system as it operates?"

(II) *Critical harnessing of social complexity:* This category consists in the individual asking herself, "How could 'our' society become a 'better' one, so that 'we' are freer than before?" It diffracts into two further subcategories:

　　i. *CA-critical harnessing:* Given the features (*a*) and (*b*) of *S*, take courses of action which will make use of (*a*) and (*b*) to pursue a *reflected* and *intra-systemically coherent* idea of freedom, i.e., an idea which the overall logic inherent in (*a*) and (*b*), as well as your location on the graph, does not spontaneously attribute to you but which remains consistent with upholding structural features (*a*) and (*b*).

　　ii. *OS-critical harnessing:* Take courses of action which use (*a*) and (*b*) to pursue a *reflected* and *extra-systemically coherent* idea of freedom, i.e., an idea of freedom whose coherent pursuit requires modifying the overall logic inherent in (*a*) and (*b*), i.e., requires modification of (*a*) and/or of (*b*).

Category I corresponds essentially to all purely spontaneous activity in which individuals in the industrial market society seek to "make the best of it," possibly—if they can afford it—with the help of professional consultants for greater efficiency, outsourcing, rationalization, better employability and flexibility, and so on. Category II*i* covers various instances of "changes of views" in the industrial market society, as when a unionized workers chooses to become a Republican militant for corporate tax cuts, or when a doctor converts to social medicine even thought his cuts her income by 80%, or when a former CEO finds he must become Minister of Social Affairs to bolster social-democratic tax policies and to introduce a Tobin tax on financial transactions. Finally, category II*ii* covers all "revolutionary" or "deeply reformist" options which are preoccupied with modifying the basic patterns of social interaction—say, replacing hierarchy with participation *à la* Hirschman (1970), or substituting decentralized planning for markets *à la* Albert (2003)—and/or the basic values and goals harbored by individuals.

　　It would be a gross misperception to immediately jump on category II*ii* as merely embodying the good old "specter of Stalinism" in which changes in interaction patterns and in "mentalities"

are more or less hammered into people at the behest of a smug but brutal intelligentsia. As all educational reformers know (see e.g. Atlan [1991]), efficiently changing mentalities or ideas will frequently involve a long, slow process of diffusion extending over generations, and always in danger of ending up where one had not intended it to. In fact, since they are concerned with *harnessing* the *prevailing* social complexity, all three above categories of emancipation are by definition concerned with *intra-systemic* interventions—and this seems to leave room for three broad categories of action in each case:

a) *Engineered criticality (lever-point):* Obtain a position of leadership or influence and act so as to induce the (sub-)system to self-organize into a critical threshold leading into a phase transition.

b) *Discursive dissemination or "contamination" (voice):* Obtain means and/or positions necessary (if any) for the dissemination of your ideas about interaction patterns and/or about individual and/or collective goals and motivation (either to defend the existing ones against attack or to promote non-existing ones).

c) *Intentional disconnection (exit):* Voluntarily cut off the directed flows to your node or to a group of nodes relevant to you on the social (sub-)network, and attempt to either reconnect into another (sub-)system or to make the disconnected group of nodes self-sufficient.

Category *i* covers many phenomena ranging from the acquisition of political mandates in order to push an agenda to the emergence of a "revolutionary avant-garde." Category *ii* covers both contrarian initiatives such as denigration campaigns and official opposition and apologetic initiatives such as propaganda and re-election campaigns. Finally, category *iii* is concerned mainly with the creation of what have become known as "intentional communities," but also with phenomena such as collective resignation of workers in order to create their own self-managed firm with new intermediaries and customers. The combination of emancipation types I, II*i* and II*ii* and of action modes *a*, *b* and *c* generates a three-by-three matrix which, although perhaps not exhaustive, seems to be a sufficient starting point for our present purposes.

1.3. The critically emancipated society as a specific issue in social problem-solving

Given the stress I placed from the outset on emancipation through critical reflection on and in complex social systems, I will mainly focus on categories II*i* and II*ii* as well as *a* and *b*. A critically—as opposed to merely opportunistically—emancipated society is one in which people seek to pursue CA- or OS-critical action and seek to do this mainly through the search for lever-points and the practice of voice, as opposed to mere exit.

In that sense, critical emancipation is a specific sub-variety of what Lindblom and Cohen (1979) have termed *social problem-solving*:

By social problem solving, we mean processes that are thought to eventuate in outcomes that by some standard are an improvement on the previously existing situation, or are presumed to so eventuate, or are conceived of as offering some possibility to so eventuate. We do not limit the term to processes that achieve ideal or even satisfactory outcomes; and in that light, "problem-attacking" is more accurate a term than "problem-solving." Nor do we limit the term to the intellectual processes through which people grapple with problems. Coin tossing is also a problem-solving activity. Some students of problem solving hold that "solve" implies understanding, as in solving a mathematical problem. For us, "solve" does not require an understanding of "the problem" but only an outcome, as when coin tossing solves a problem of whether to turn left or right at an unfamiliar, unmarked road junction. (Lindblom and Cohen [1979: 4])

This considerably broadens and de-intellectualizes Karl Popper's (1962) famous and foundational "fourth thesis" concerning the structure of social science, namely: "Knowledge does not start from perceptions or observations or the collection of data or facts; it starts, rather, from *problems*" (Popper [1962: 65]). Contrary to Lindblom and Cohen, Popper does require an understanding of the problem to be solved—essentially, we have to be able to model the problem in a hypothetical form conducive to strictly logical deductive operations—and he imposes on the derived solution the imperative of what he calls "criticism." As is well-known, he means by this simply that "all criticism consists in attempts at refutation" and "the logical instrument of criticism—the logical contradiction—is objective" (Popper [1962: 67]).

One of the possible—and, the Frankfurt School Marxists would claim, in our alienated industrial market societies, very likely—implications of the Popperian imperative of critically-rational problem-solving is that the problem of *critical* social emancipation either would turn out to be a problem of very limited scope or would be seen as based on empirically false assumptions about the motivations of individuals. Where, the logical positivist would ask, do you find today your empirical correspondents of individual "aspirations to emancipation"? You will at best be able to gather some polling or questionnaire data on people's aspirations to a better life for themselves (category I), implying either a view of having to go it alone (category *c*) or a mildly

vocal call (category *b*) for certain intra-system changes (category II*i*): more taxes or less taxes, new regulations to bolster or to limit the freedom of market actors, and so on. Certainly, you are unlikely to find large portions of the population harboring alternative views of the very structure of society (category II*ii*) and believing in the acquisition of lever-point positions (category *a*) other than through the democratic election of representatives who promise to manage industrial market society from within its self-imposed limits. Horkheimer and Adorno would not dispute these *prima faciae* observations, but they would deny them the absolute status given them by falsificationism. In his reaction to Popper's conference on the structure of social science, and in reaction also the spread of pragmatist and "common-sense" philosophies which also impelled Horkheimer's (1946) critical efforts, Adorno seeks to re-open the scope for nontrivial aspirations and thus for making OS- as well as CA-critical emancipation part of social science. He does so by contesting the received demarcation between psychology, which deals with observable "facts" (types, etc.) and sociology, which has been largely co-opted by psychology but ought to investigate *non-observable factuality*, i.e., virtual tendencies which fail to come to the fore precisely because the individuals who harbor them are embedded in a social objectivity which denies them. Sharing Popper's anti-psychologism but deducing from that common rejection some radically different implications, Adorno claims that a too precocious anti-psychological stance actually fails to see the momentary truth of what it too quickly wants to brush under the carpet:

The separation between man and the social environment [is] a fact given in the current landscape of science whose hypostasis Popper fundamentally rejects. The subjects which psychology purports to examine are not merely, as the usage has it, influenced by society, they are intimately formed by it. The substrate of "man in itself," which supposedly would be facing the environment, would remain an empty abstraction. Conversely, the socially active environment is a product—albeit an indirect and ill-recognizable one—of men, of organized society. Nevertheless, psychology cannot be considered the bedrock discipline of social science. (…) Society is a general process in which men, surrounded and steered and formed by objectivity, nevertheless react back on that objectivity (…). Only those who are able to think society as something other than the existing one can come to perceive it as a problem, in Popper's vocabulary; only through that which it is not can society reveal itself as what it is—and that is indeed what should be most at stake in sociology if it did not, as is the case in most of its projects, remain content with catering to the aims of public and private administration. (Adorno [1962 : 563-564])

When we take "sociology" as a near-synonym to social science in general, this statement amounts to submitting to critical scrutiny the very criterion which Popperians use for their brand of criticism, namely the separation of facts from theories, seeing theories as representational devices to be deductively tested against "the facts." The Frankfurt School's point is that the way any scientist is able go about collecting, interpreting and even eliciting "the facts" from empirical individuals is inextricably bound up with the way in which overarching social mechanisms

(which, Adorno acknowledges, emerge from interactions of individuals) form these individuals' modes of perceiving themselves and their place in the social world. Thus, Raymond Geuss's methodological position, also to be found in the work of Raymond Boudon (2003) and others, that "I want to keep the notion of 'want/desire' closely tied to avowal and behavior" (Geuss [1981: 47]) simply cannot be upheld: "The renunciation of sociology to a critical theory of society stems from resignation: we no longer have the nerve to think the Whole because we despair of ever being able to change it" (Adorno [1962: 565]).

But this is the crucial point: today's theories of complex adaptive systems *do* "think the Whole" in very explicit terms indeed, and the general idea associated to this strand of science that lever points and other tools for harnessing social complexity may be available (though not yet in a fully systematic way; see Holland [1995: 161-172]) contains none of the demarcations we have just encountered between observable facts and imagined aspirations, or between the agents' avowed goals and the modeler's abusive projections. Thus, *the general issue of "harnessing complexity" prescribes no a priori skepticism against either CA-critical or OS-critical harnessing.*

2. The phenomenology of critical instrumental rationality: From "spontaneous consciousness" to "conscious spontaneity"

2.1. Hayekian Critical Theory: What balance of interactive spontaneity and reflective consciousness?

There is no way of circumventing the social thought of Friedrich von Hayek when one is investigating the harnessing of complexity. Indeed, as Dupuy [1992a, 1992b] has persuasively shown, such circumvention would be downright harmful because for all his controversial traits Hayek was and remains one of the most profound philosophers of liberalism. One of his chief insights, encapsulated in the intuition that social phenomena are "the result of human action but not of human design" (Hayek [1967]), is that

Most of the knowledge on which we rely in the pursuit of our ends is the unintended by-product of others exploring the world in different directions from those we pursue ourselves because they are impelled by different aims; it would never have become available to us if only those ends were pursued which we regarded desirable. (…) *Men can be allowed to act on their own knowledge and for their own purposes only if the reward they obtain is dependent in part on circumstances which they can neither control nor foresee.* (Hayek [1976: 111 and 120])

The members of Frankfurt School have harsh words for this insight—not because it is factually inexact (in industrial market society, as well as in earlier social formations, it is true that most individuals feel unable to influence the aggregate *faits sociaux*) but because they feel Hayek is using it in an apologetic and somewhat naturalistic way to legitimize capitalist alienation and exploitation, as well as a "rule of law" which may be little more than the legal arm of the most powerful agents in the system:

The insights into society as a totality (…) imply that all moments which concur to this totality, and which are in no way reducible to one another, must be integrated into knowledge; the latter must let itself be terrorized by the scientific division of labor. The priority of the social entity over the human individual [in holist doctrines] can be explained by that very impotence of the individual in the face of society which was Durkheim's main criterion for *faits sociaux* (…) The *reductio ad hominem* which inspires all of critical enlightenment has as its subject matter that human being which a self-mastering society has yet to produce. (Adorno [1962: 563 and 565])

Clearly, if we want to avoid that the "self-mastering society" be of the totalitarian variety—and Horkheimer himself, who laments "[t]he takeover of what belongs to the individual into the state's keeping (Horkheimer [1937b: 248]), would certainly concur—that self-*mastery* must be of a self-*organizing* kind. In other words, collective mastery of social phenomena must be an emergent property of the social system, hence of the multifarious interactions of the individuals. This is indeed one Hayekian insight which can no longer be rejected; the crucial issue is whether Hayek is not ill-guided when, in addition to self-organizing interaction, he postulates the necessity for the "knowledge" used by individuals to be merely the "very important but unorganized knowledge (…) of the particular circumstances of time and place" (Hayek [1945a: 80]). I claim he is indeed misguided in that respect. Philip Ball puts it very cogently:

Hayek is right at least to believe that there *are* spontaneous forces of society. (…) No doubt some of these forces are what we might call exogenous: they arise from outside, from technological change for example (such as mechanization, birth control, information technology), or from the environment (drought and famine, climate change). But many of the forces that act on society are the collective consequence of interactions between one person and another, whether that be in conducting trade, spreading new fashions, waging war, getting married or avoiding collisions. (Ball [2004: 582])

Apart from birth control which emerged in part from the emancipatory aspirations of women embodied in feminist social movements (but also, Frankfurt School members might add, from the need for budding industries and bureaucracies in the 1960s to enroll a new workforce eager to enter the labor market), all the artifacts and cultural features mentioned by Ball are embodied in our spontaneous perceptions of ourselves, others and the collectivity within the industrial market society. They are, Horkheimer would say, "*bourgeois*" traits and objects. As such, they may well

be the result of Hayekian "cultural evolution," but they have become an integral part of most individuals' unreflective interactions. I call these elements the individuals' *spontaneous consciousness*—a set of values, conventions, concepts, methods of investigation and production, etc., which form the very fabric of social interaction in industrial market societies. Spontaneous consciousness is, if you like, the backbone of the evolved social system's directed graph(s) defining its interaction pattern(s) and the motivations that generate flows between nodes of the graph(s).

Let us now pause for a second, however, and reflect on the following question: what if one of the things individuals did in those "interactions between one person and another" was to communicate and exchange—through direct speech acts but also through indirect ones such as books, newspapers, films, music, works of art, television documentaries, and so on—models and theories of what a better society, i.e., improved interaction patterns and/or motivations, might or should or could or ought to be like? Hayek himself very early in his career did allow, following "the great writers of the eighteenth century," that an individual's "self" might well include "anything for which people in fact did care" (Hayek [1945b: 13]). *So why should "what I care for" not include my own aspirations to a better society—whether CA-critical or OS-critical?* This is certainly what Horkheimer has in mind when he writes forcefully that

[The] difference in the existence of man and society is an expression of the cleavage which has up to now affected the historical forms of social life. The existence of society has either been founded directly on oppression or been *the blind outcome of conflicting forces*, but in any event *not the result of conscious spontaneity on the part of free individuals*. Therefore the meaning of "activity" and "passivity" changes according as these concepts are applied to society or to individual. In the bourgeois economic mode *the activity of society is blind and concrete, that of individuals abstract and conscious*. (Horkheimer [1937b: 200], my italics)

In other words, while social phenomena are blind (they emerge) but concrete (they constrain), individuals' activities are conscious (they apply culturally inherited means to pursue their culturally inherited ends) but abstract (they base their actions only on very localized knowledge of "the particular circumstances of time and place"). This, according to Horkheimer, is the epitome of an alienated situation—and he believes it can only be resolved if individuals move from what I have called interactive spontaneity to what he calls conscious spontaneit*y*, or what I would call *reflective consciousness*. The terminology matters little; what counts is the underlying issue: *how does a complex interactive system function when its constituents—here, human beings—interact on the basis of critical-reflective judgments about that system itself, and not just on the basis of whatever local knowledge and information is available at their particular node in*

the social network? Dealing with this problem is, in essence, the objective of what I call Hayekian Critical Theory (see Arnsperger [2004c]).

Hayek's first-hand reply to the question would probably be that it is a wrong question to ask, because if individuals become social critics they will attempt to influence each other into acting on overall judgments and into adopting a bird's eye view of the system—a *theory*, i.e., a contemplative vision—which no amount of knowledge can provide. But of course, this cannot be right because, as any Hayekian is fully aware, already in the game of catallaxy (unless they are playing a Walrasian equilibrium, which Hayek explicitly rejects; see Hayek [1945a: 91]) individuals directly or indirectly influence each other, including in transmitting through prices and rationing information about what objectives it is right to pursue. Thus, at the very least, each Hayekian individual must have his own rationalization of why she is playing such a game, and this cannot be achieved simply through blindly accepting that whatever "local" rules and conventions have been transmitted to her by the past generation of market actors in her family and community is "what's to be done"—the rationalization need not be a fully articulated theory, it can be merely a set of normative principles inherited along with the practices and cursorily legitimizing "what's to be done." Now, these principles certainly will not be apt to *predict or foresee* any of the outcomes of the complex market system, but they will be apt to embrace in a few (half-fatalistic, half-normative) statements the whole of the forces of market society—"competition," "making a profit," etc.—so as to ground the individual's actions. Why, then if not by a misguided confusion between comprehension as prediction and comprehension as theory, does Hayek insist that "the part of our social order which can *or ought* to be made a conscious product of human reason is only a small part of all the forces of society" (Hayek [1945b: 22])?

The fact is that if we are all participating in market forces on the basis of a simple creed—a simple, highly simplified descriptive and/or normative theory of the market society—the whole of our social order *will* be made a conscious part of our reason, if only because we will thereby rescind from objecting to certain outcomes or criticizing certain mechanisms. If that is so, then whatever simple theory helps us achieve this state of quietude or acceptance is one which cuts off categories II*i* and II*ii* above, and safely parks us within the arena of category I. Being aware of this necessarily prompts the question, Why? And if Hayek replies that this "ought to" be so because such ignorance is needed for the market order to function, we can only bend his own

point back on him: How do you know, if not on the basis of a very specific theory of the Great Society, i.e., the Market Society? If Hayek next replies that his is not a *theory* of the market society but a mere reflection on our *de facto* cultural heritage, we can again counter that *another* part of that *same* heritage is that of critical rationalism, and that calling it *ex ante* a source of "abuse of reason" (Hayek [1952]) only begs the question.

There is no honest way around it: a Hayekian society has to allow for the whole *Shebang*—individuals, if they are to adhere to the game of catallaxy, have to adhere to it *consciously* because their *critical faculties* are part and parcel of the heritage which liberalism itself has bestowed upon them.

2.2.   *The individual as social critic: Critical instrumental rationality is not (just) social science*

One of the key contributions of the Frankfurt School is to have so profoundly renewed the old Marxist theme of "theory and praxis" that we can truly speak of an *individualistic model of critical instrumental rationality*. In Marxist fashion, Horkheimer and Adorno reject the theory-praxis division and claim at least two things: theory-building and theory-changing is itself an activity in society, with its more or less acceptable norms and procedures; moreover, a *critical* theory is thought to equip its bearer not only with descriptive concepts or with demystifying ones, but also with a capacity to reflect on *means* to reach his *end*, namely the transformation of society from its current to a better state. Geuss has provided a remarkable and exhaustive characterization:

A typical critical theory will be composed of three main constituent parts:
- (A) A part which shows that a transition from the present state of society (the "initial state" of the process of emancipation) to some proposed final state is *"objectively"* or "theoretically" possible, i.e. which shows:
  1) that the proposed final state is inherently possible (…);
  2) that it is possible to transform the present state into the proposed final state (…).
- (B) A part which shows that the transition from the present state to the proposed final state is *"practically necessary,"* i.e. that
  1) the present state is one of reflectively unacceptable frustration, bondage, and illusion (…);
  2) the proposed final state will be one which will lack the illusions and unnecessary coercion and frustration of the present state (…).
- (C) A part which asserts that the transition from the present state to the proposed final state can come about only if the agents adopt the critical theory as their "self-consciousness" and act on it. (Geuss [1981: 76])

What can impel an individual to "adopt [a] critical theory as [her] 'self-consciousness' and [to] act on it"? Understanding this requires that we delve more in detail into what Adorno earlier called the "critical impulse."

The philosophical notion of criticism refers essentially to a specific cognitive move by which, as it were, the "organ" of knowledge bends back on itself and examines itself. *Criticism as self-examination* is, some would say, as old as philosophy—in the Socratic tradition, it *is* philosophy, and one of the great intellectual conquests of the Enlightenment was to merge this subject-oriented Socratic tradition ("Know thyself") with the growing awareness that there is a natural and social world "out there" that feels more like an object of knowledge. At least from Kant onwards, the distinction between the knowing subject and the knowable object(s) has become increasingly blurred, to the point where in cognitive science the alleged subject ("I," the brain) becomes his own object while, in critical sociology, the alleged object ("society") becomes a specific kind of active subject. Now, it turns out—and this is crucial for our present endeavor—that this blurring of the borders between subject and object of knowledge is closely connected with the historical emergence of invisible-hand explanations and the critique of individual impotence which these explanations are apt to elicit. Horkheimer offers this striking parallel between Kant's incomplete transcendentalism and the Hayekian attempt to posit individuals as ignorant of the overall social forces shaping their lives:

The difficulty and obscurity which, by Kant's own admission, mark the sections [of his book that deal with] the deduction and schematism of the pure concepts of understanding may be connected with the fact that Kant imagines the supra-individual activity, of which the individual is unaware, only in the idealist form of a consciousness-in-itself, that is a purely intellectual source [i.e., the transcendental subject]. In accordance with the theoretical vision available in his day, he does not see reality as product of a society's work, work which taken as a whole is chaotic, but at the individual level is purposeful. (…) The activity of society thus appears to be a transcendental power, that is, the sum-total of spiritual factors. However, Kant's claim that its reality is sunk in obscurity, that is, that it is irrational despite all its rationality, is not without its kernel of truth. The bourgeois type of economy, despite all the ingenuity of the competing individuals within it, is not governed by any plan; it is not consciously directed to a general goal; the life of society as a whole proceeds from this economy only at the cost of excessive friction, in a stunted form, and almost, as it were, accidentally. (Horkheimer [1937b: 203])

For Horkheimer, the aim of critical social science is to complete this incomplete project by integrating into rationality the reflection on the overall "frictions" of industrial market society—frictions which *in the individual's experience* become reified into certain economic mechanisms that appear fateful. To overcome this "tragic knowledge" (as Christoph Menke [1996] has aptly termed it) and to move towards a constructive critique is the very aim of the Frankfurt School's theoretical efforts.

To a Hayekian mind these efforts must ultimately be quixotic because the very "fatefulness" of the aggregate social phenomena when view from the individual's perspective is precisely what allows the *autonomy of the social* so dearly required in a free society (Dupuy [1992a]).

Nevertheless, profound dissatisfaction or even revolt in the face of unmastered social forces is what drives the "critical impulse" in the first place. Hayek's deep error—and here I side decidedly with Horkheimer and Adorno—is to have confused the level of scientific explanation with the level of lived experience. As we shall see shortly, the *primacy of experience* that inhabits some strands of contemporary cognitive science forbids a wholesale denigration of the concrete, embodied experience of dissatisfaction or revolt, all the more so because that very experience may itself be a product and a trigger of social interactions, and because some experiences of *collective action* do show that allegedly fateful mechanisms *can* be modified, sometimes even deeply. So, whether her experience is scientifically rationalizable or not, the individual is *de facto* a social critic as soon as she reflects on her socially lived experiences and integrates them into either an opinion or a theory as to what is currently wrong and what should be done to correct it.

Classical accounts of "the social critic," such as those of Michael Walzer (1987, 1988) and Edward Said (1994), share with the Frankfurt School a bias towards the critical *intellectual* as opposed to the critical *citizen*. That bias may be understandable in the sense that (as this paper itself evidences) we often see intellectuals—writers, academics, and so on—and intellectuals often see themselves, as a privileged channel through which critical opinions and theories about existing society can be elaborated and disseminated. However, this role as a kind of crucial "node" in the social network does not get us very far if we cannot account for the fact that some or all other "nodes" actually *receive and assimilate* the flows of ideas which originate in the privileged "nodes." Thus, tendentially, there is no essential difference between the category of the critic and the category of the non-critic—it is a passing, evanescent difference attesting to one of the Enlightenments most enduring postulates: the fundamentally equal capacity of all individuals to develop and cultivate critical reasoning. "What ordinary people do to achieve coherence does not greatly differ in main outline from what scientists do in their scientific work" (Lindblom [1990: 40]). Therefore, we have to view critical instrumental rationality as being itself both the result and the stimulator of social interconnectedness. Walzer use social-network vocabulary when he describes how the common man becomes part of "the company of critics":

Disappointment is one of the most common motives for criticism. We have an idea about how institutions ought to function or how people ought to behave. And then something happens, the authorities act with the usual brutality; or something doesn't happen, the people are passive and indifferent; and we feel ourselves thrust into the company of social critics. It takes some further motivation, though, actually to join the company and stick to the critical enterprise. Disappointment isn't enough. Nor does a disinterested desire for the well-being of humanity seem a sufficient motive. A moral tie to the agents or the victims of brutality and indifference is more likely to serve. We

feel responsible for, we identify with particular men and women. Injustice is done in my name, or it is done to my people, and I must speak out against it. Now criticism follows from connection. (Walzer [1988: 22-23])

The fundamental thrust of everyday-life social criticism, or what Walzer calls "connected criticism," is therefore a subtle and individually specific mixture of solidaristic identification (Arnsperger and Varoufakis [2003]) and emotional involvement (Livet [2003]), triggered by the concrete experience of disappointment. None of these categories even get near to a Popperian ideal of scientificity: a critical theory as set out above by Geuss will quite often be unfalsifiable, or make use of "utopian" concepts built up in protest of current social conditions but unobservable given these conditions—but why should that deter the individual's cultivation of critical instrumental rationality, i.e., her continuous reflection on what is wrong today, on what would be better, and on what would be the best means to effect an improvement? *If it is part of the current contradictions or "frictions" of industrial market society that individuals elaborate unscientific and utopian theories, on what "scientific" grounds can anyone tell them they are "wrong"?* Geuss puts it with characteristic clarity:

Agents can act in ways that are more or less enlightened; the freedom of communication and discussion they enjoy and their freedom to form and acquire beliefs and preferences is a matter of degree; agents can be more or less reflective. To what extent a critical theory is enlightening and emancipatory may then equally be a matter of degree. If rational argumentation can lead to the conclusion that a critical theory represents the most advanced position of consciousness available to us in our given historical situation, why the obsession with whether or not we may call it "true"? This is not a form of the relativism the members of the Frankfurt School rightly reject. If it is closer to Adorno's historicism than to Habermas's recent views about the "idea speech situation," that seems to me to be an advantage: the critical theory is better off without the transcendental baggage. If a critical theory is not a true "scientific" theory, not a part of empirical social *science* strictly so called, we might think of it as part of the wider enterprise of social theory or social philosophy. Not all empirical social inquiry must have the structure of critical theory, but the construction of an empirically informed critical theory of society might be a legitimate and rational human aspiration. (Geuss [1981: 94-95])

This strongly echoes Charles Taylor's important point to the effect that theories in general do not simply have explanatory uses, but also serve "to define common understandings, and hence to sustain or reform political practices, as well as serving on an individual level to help people orient themselves" (Taylor [1983: 107]). In Taylor's parlance, we can call this function of theory the *self-definitory* function. Self-defining theories, among which are the critical theories held by "connected critics," need not be scientific in Popper's sense—that is Geuss's main point. Now, given what we said before, we should rather phrase his last sentence as follows: "...the construction of an empirically informed critical theory of society *is* a legitimate and rational aspiration *of each individual*." And if this is so, we know that two key aspects of this construction are (*a*) the individual's concrete experience of what to criticize and (*b*) the interactive context

within which this experience arises and becomes an object of reflection. Let us now turn to these aspects.

2.3.   *How empirically plausible is critical instrumental rationality? The "enacted" component of reflective consciousness*

One of the most common complaints addressed to the figure of the individual as social critic is that it is empirically implausible. People, it is often said, *are not like that*; they can't be bothered to cultivate critical instrumental rationality because they care for more immediate things in their environment—house, family, tomorrow's income, next year's vacation. Thus,

many kinds of studies of ignorance and irrationality treat them as somehow fixed, hence ignore their social causation. Citizens simply do not bother to probe—but why should they? They are interested instead in jobs and "happy families." Thus is dismissed the study of social causation of the defects (…)—that, in short, they have been deliberately or inadvertently induced or caused to probe badly rather than well often gets no more than a formal acknowledgement rather than study. (Lindblom [1990: 181-182])

What Lindblom calls probing well rather than badly means, in our present framework, cultivating and using critical instrumental rationality instead of uncritical instrumental rationality or non-instrumental (idealistic) critical rationality. And that framework, Walzer's following formulation can put us on the right track:

The tasks of the critic (…) are (…) to question relentlessly the platitudes and myths of his society and to express the aspirations of his people [as he sees them, including his own; see Geuss (1981: 45-54)]. The second of these isn't possible, obviously, unless the people actually have aspirations that reach beyond or clash with the social order in which they live. If the masses are entirely satisfied (…), then, the critic is left with the thankless task of criticizing their satisfaction, which he hopes to replace with his own unfulfilled aspirations, the products of detachment and solitary reflection. But why should anyone accept the replacement? And what can the critic do when it is refused? These are difficult questions for the critic-at-large, but they need not call into doubt the critical enterprise itself. Imagine a critic who isn't detached, free-floating, or alienated. He understands himself instead as a social being, "a man of a certain region, a certain class, and a certain time" (Silone), critical of "his own world" (Bourne). His values, even his universal values, are first of all the values of a particular person, and they are shared with a particular set of other people: "average values" (Camus). (…) Though he starts with himself, he speaks in the first person plural. This is what we value and want, he says, and don't yet have. This is how we mean to live and don't yet live. We criticize our society just as we criticize our friends, on the assumption that the terms of the critique, the moral references, are common.

Now, the question as to how close critical rationality should be to either "detachment" or "average values" cannot be settled as easily as Walzer suggests, lest we want to throw social criticism directly back into the arms of Hayek's narrowly construed cultural evolution. However, what Walzer does convey convincingly is that the cultivation of critical rationality is an embodied experience, a localized attempt to reach for universality within a social network. This means, among other things, that the critical individual's social world is not simply something that "informs" her brain of objective facts merely to be "processed" in order to obtain a neat

"representation" of the world, its defects and its yet unexploited possibilities. Rather, without much of the "transcendental baggage" (Geuss) that Habermas would like us to carry along, we come to know our social world and ourselves at the same time by acting in it unknowingly at first—our "self" is not so much the precondition as the result of what Varela and his colleagues call *enacted knowledge*, based on the recognition that

there are many ways that the world is—indeed even many different worlds of experience—depending on the structure of the being involved and the kinds of distinctions it is able to make. And even if we restrict our attention to human cognition, there are many various ways the world can be taken to be. This nonobjectivist (and at best also nonsubjectivist) conviction is slowly growing in the study of cognition. (…) We propose as a name the term *enactive* to emphasize the growing conviction that cognition is not the representation of a pregiven world by a pregiven mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs. (Varela, Thompson and Rosch [1991: 9])

The history of which the authors speak is, as we saw earlier, the sequence of an individual's "structural couplings" with his environment, on the background of that individual's "operational closure," i.e., the internal coherence which makes the brain able to string together the various crises and ruptures of microidentities and microworlds as repeated couplings and decouplings occur in sequence. As phenomenology has long recognized—and the Levinasian critique of Husserl notwithstanding—there is no fundamental ontological distinction between a structural coupling with an object and a structural coupling with another human being, even though the content of the experiences and the nature of the flows of content and information will differ greatly… Thus, human-to-human interaction is an integral part of the "perturbating" or "irritating" experiences that build up each individual's enacted/embodied knowledge of the social world. The authors suggest that this conception of cognition also modifies our conception of reflection:

What we are suggesting is a change in the nature of reflection from an abstract, disembodied activity to an embodied (mindful), open-ended reflection. By *embodied*, we mean reflection in which body and mind have been brought together. What this formulation intends to convey is that reflection is not just *on* experience, but reflection *is* a form of experience itself—and that reflective form of experience can be performed with mindfulness/awareness. When reflection is done in that way, we can cut the chain of habitual thought patterns and preconceptions such that it can be an open-ended reflection, open to possibilities other than those contained in one's current representation of the life space. (Varela, Thompson and Rosch [1991 : 27])

Thus, the precarious individual act of striking the balance between Hayekian "spontaneous consciousness" and Horkheimerian "conscious spontaneity," so as to cultivate critical instrumental rationality, is in itself a component of the individual's social experience. In their all too brief discussion of social-exchange theory (Varela, Thompson and Rosch [1991: 245-246]),

the authors indicate that solidaristic identification and emotional involvement are key elements in the *opening-up of ingrained representations of our social contexts and our goals to be pursued within those contexts.* Thus, enacted knowledge of the dissatisfactions or revolts around us, and embodied participation in the interactions that yield our self-definitory social theories, is an essential empirical part of the genesis of critical instrumental rationality.

Brought together in this fashion, critical social theory and enactive cognitivism offer a fairly radical renewal of cognitive science, allowing for a stimulating integration of social reflexivity into interactive spontaneity, and of social spontaneity into interactive reflection. In fact, the two cannot be separated but this implies in no way that we need to relinquish the basic tenets of Hayekian methodological individualism. On the contrary, as we shall see presently, recasting the interplay between individual rationality and emergent social wholes within the Frankfurt School framework makes for a progressive and dynamic research program.

3.   Neither transcendental nor dialectical reason: Toward a non-holistic theory of self-criticizing social systems

3.1.   "Esprit critique" as an emergent property: How a complex multiplicity of social critics can achieve operational closure

Interaction between critical instrumental rationalists can lead to paradoxical situations because all *n* of them interact within a single prevailing system *S* while having as one of their objects of interaction, and perhaps even as the single objective underlying all their everyday interactions, opinions and theories about what their interactions should be like in any of several ideal systems S^*_k, where *k* is an element of $\{1,\ldots,K\}$, with *K* the number of groups having distinct ideals. Thus, for instance, several groups may in an industrial market society enter into competition on the "market" for anti-competition ideals—Trotskyists, left-wing anarchists, moderate advocates of a self-managed economy, neo-romantics, etc., may all be scrambling for commercial space, well-selling publishers, and prime-time media niches so as to get their ideas to the largest possible audience. Feminists militated for decades within a man-made world in which they had to play along while at the same time being aware of how much dislike they felt for the prevailing patriarchal rules and practices. More generally, for any social mechanism *m*(*S*) inherent in the structure of the prevailing system *S*, i.e., determined by its directed graph of social relations and the associated goals and motivations of individuals, there will be a set of groups using *m*(*S*) while

27

promoting anti-$m(S)$ ideals, hence ideas which judge S itself to be structurally undesirable. The apparent paradox vanishes, however, once we renounce our implicit ideas about "social stability" and "agreement" and come to understand such a system S as a *self-criticizing social system*.

Now, this notion is emphatically *not* a holistic one—in other words, the "self" as used here is not a subject, but merely a reflexive pronoun designating a redoubling of the same entity, in the same way as we call an organ or a tissue "self-repairing" without in any way lending it the qualities of a subject. Indeed, a system whose members use its inherent mechanisms to criticize those very mechanisms and hence the system itself is a subjectless whole *par excellence*, in the purest Hayekian tradition. Being so, it also stays clear of the implicit (and often explicit) functionalism and quasi-organicism of the great systems theorist Niklas Luhmann. Nevertheless, Luhmann's intuitions cannot be brushed aside too quickly. Here is what he has to say about the role of criticism in society:

With the increase in the complexity of society, all possibilities get used more heavily and in a more functionally specific way. It becomes less alarming that at the level of interaction contacts are interrupted, Christmas greetings are not answered, marriages dissolve, firms crumble. But that difference is connected with a more or less established equilibrium between what stops and what begins anew. Moreover, the structural instructions for reproduction become more specific, which makes them more fragile and more quickly obsolete. Both ways of reacting to higher complexity harbor their own conditions and their own problems. They seem not to be sufficient by themselves. Therefore, in parallel, the society's immune system gets reinforced, too. It does not consist merely in a negative copy of the structures, nor merely in a "critical" awareness vis-à-vis the available, but rather in specific, distinctive forms of the continuation of communication—in forms such that, for instance, situations generated through struggle and victory vary so much as to constantly make subsequent normalizations possible. Within the framework of this selective formation of dissent and conflict, the increased strength of positions of legally permitted refusal, as well as the articulation of unrest, critique and protest in the guise of social movements, have both become more significant. In traditional representations of social history, they are usually played out one against the other: on the one side, the political and economic complex of modern capitalism; on the other side, the correspondingly stimulated whole of social movements. It would be more theoretically sound to distinguish between the structure of expectations and the immune system. One would then certainly realize that, compared with all its historical ancestors, modern society has destabilized its structures and has greatly increased the ability to say "No." It may then be less important to know whether this "No" gets articulated more from positions of legal strength or more in the context of social movements. Nowadays there is a tendency to try and reconcile these two perspectives in the figure of "civil disobedience." Whatever the case may be, we have to ask ourselves how from that situation we can recapture the still necessary "Yes" to society. (Luhmann [1975 : 549-550])

This last question is no problem for self-criticizing social systems: all those who want to say No say it while provisionally saying Yes to the structures in which they are saying No. The seeming contradiction inherent in such a situation is simply rooted in the Enlightenment tradition itself, and one of its versions is the one highlighted by Horkheimer in his discussion of Kant: we precipitate the emergence of social forces we cannot foresee while at the same time lamenting them and attempting to overcome their detrimental effects through our "critical" participation in them. So Luhmann's downplaying of critical awareness seems to me a rather ideological

stance—although one which, as I understand his approach, is motivated in large part by his earnest faith in functionalist systems theory. Thus, in a system whose internal coherence requires "immunity," rather than mere homeostatic re-stabilization, saying "no" can only make sense if it actually reinforces one's fundamental "Yes."

This neglects, however, a category of attitudes which is crucial for the Frankfurt School, and which Luhmann seems to minimize—namely, the attitudes that can be grouped under "critical acceptance" of norms. As Horkheimer puts it, "the critical acceptance of the categories which rule social life contains simultaneously their condemnation" (Horkheimer [1937b: 208]). Condemning a rule, or a whole social structure, does not leave much room for contributing to its "immunity"—what would it have meant for feminists to condemn patriarchal structures while contributing (even unwittingly) to their immunization? My fear is that the Luhmannian answer will be quite unambiguous: *they* would not have contributed to their immunization, but the male part of society would have reacted in self-defense—much like antibodies rushing to the site of infection—and defend themselves they did. Fair enough; but what does this tell us, other than that these heroic women's critical instrumental rationality led them to incur large sacrifices to their immediate interests for self-preservation out of commitment to a cause (see e.g. Sen [1977], Arnsperger and Varoufakis [2003])? In many cases, therefore, saying "Yes" to the currently prevailing social arrangement may mean nothing more than staying in it because one has nowhere else to go, so that "here and now" is the only place to root one's critically motivated actions.

Still, Luhmann's basic preoccupation with the need to account for the social system's operational closure is certainly warranted, even if his more conservative-sounding solutions are not. Unless we are ready, in the face of a proliferation of critical instrumental rationalities, to embrace the extreme microreductionist position that "there is no such thing as a society," how does a self-criticizing complex system achieve this closure? In other words, how does the structure of relations between the individuals in such a system lend the system its internal coherence? The key is to be found in our earlier characterization of the three types of harnessing—opportunistic, CA-critical and OS-critical. The very term "harnessing" means that one starts from where one is located; and notice that indeed all our three modes of harnessing, whether relying on a notion of freedom which is unreflected (opportunistic harnessing), reflected and intra-systemically coherent (CA-critical harnessing), or reflected and extra-systemically coherent (OS-harnessing), contain the expression "make use of (*a*) and (*b*)," which refers to the

agent's acceptance of the overall social structure. This acceptance can be uncritical as in categories I and in some instances of category II*i*, or critical as in other instances of category II*i* and in category II*ii*.

What is Hayekian about this setup—and crucially so—is that while in the narrow view of markets as being sufficient for the optimal use of knowledge in society the only emergent property of the social system is the vector of market prices (or, more likely, the segmented distribution of price vectors depending on relative search intensities and other localized parameters), in this model with instrumentally critical rationalists *something much broader and significant than prices* emerges from their interactions and will be seen to provide a paradoxical but real operational closure. What is it? Well, to put ourselves on the right track, we can start by noticing this very Hayekian-sounding passage from Lindblom:

> It seems reasonably clear (…) that probing of social problems requires the participation of vast numbers of people, most of whom bring significant though greatly limited competence to their inquiries, and many of whom bring educated and experienced competences to the task. Call such a state of affairs multiplism. (Lindblom [1990 : 233])

Of course, the content needs to be adapted to our present framework where agents interact on the basis of a specific form of "competence" and give a very specific form to their "probing." Nevertheless, the overall direction of the statement is clear: multiplism is called for because "probing of social problems *requires the participation of vast numbers of people*"—but why does it *require* this? Instead of becoming suspicious of a resurgence of functionalism here, let us make an immediate parallel with Hayek's insistence that the market-price vector, also generated by the participation of vast numbers of people, is the best way of "communicating to [any agent] such further information *as he needs* to fit his decisions into the whole pattern of changes of the larger economic system" (Hayek [1945a: 84], my italics). So prices are *required* because to orient herself in the economic sphere each agent *needs* information that his immediate context does not reveal. In the same way, the participation of vast numbers of people in social problem solving *à la* Lindblom is *required* because to orient herself in her everyday social life each agent *needs* cognitive resources which her immediate context does not make available to her. And by extension, the participation of vast numbers of people in criticizing prevailing society is *required* because to orient herself in her critical endeavor each agent *needs* reflective resources which her immediate context does not make available to her.

If this is so, then what is the correspondent of the price vector? I propose to put this as the central claim of this paper:

First Proposition of Hayekian Critical Theory. *In the same way as market prices are a key emergent property of a complex adaptive system of narrowly instrumentally rational individuals, a key emergent property of a complex adaptive system of critically instrumentally rational individuals is the social system's "esprit critique." This French expression, which is more suggestive than the English "critical spirit," is not meant as a metaphor for some fuzzy or intangible spiritual entity. Just because it is not a set of numerically measurable quantities such as prices, it designates something quite definite, namely the overall "critical network structure" of the society, which allows individuals (a) to form their initial aspiration for a better society and (b) to flesh out this desire with critical opinions or theories which change through contacts with other, similarly active individuals.*

Even Hayek's idea of cultural evolution can be transferred to this framework with the appropriate adjustments: what co-evolves is not merely the rules and practices that make competitive market transactions viable, but the whole set of rules and practices (about which more below) which make "norm-regulated transactions" in ideas viable. And as this crucial part of the rule of law gets installed, we can witness more and more individuals engaging in critical reflection even on the market norms which allegedly have a transcendental status in Hayek's narrow theory of society—in other words, my present version of cultural evolution would give rise, over time, to a society of individuals who critically accept to play market games while actively seeking to subvert them, so that *the culturally evolved subsystem of the market and its associated regulations would no longer have any reason to be considered stable in all cases*. It might remain stable, but only if all or a significant majority of individuals have embraced Hayek's narrow theory of society—but why should such a market society remain forever "immune" (in Luhmann's sense) to critical experimentations on its margins and at its core, any more than communist societies can remain forever immune to critical inspirations from outside? The scope of cultural evolution is therefore significantly by the above proposition.

That proposition can, of course, hardly be called a theorem at this point. Nor is it merely a hunch or an intuition. I take it to be more of a heavily underpinned suggestion for the broadening of Hayekian social science. Let us now try to get at least some sketchy insights what the dynamic process in such a society would be.

3.2. *Neither cognitive consensus nor ideal speech situation, but unstable compromise: The complex dynamics of collective self-criticism*

John Holland has aptly described, in a nutshell, the difficulty linked to complex adaptive systems:

[Complex adaptive systems are] systems composed of interacting agents described in terms of rules. These agents adapt by changing their rules as experience accumulates. In complex adaptive systems, a major part of the

environment of any given adaptive agent consists of other adaptive agents, so that a portion of any agent's efforts at adaptation is spent adapting to other adaptive agents. This one feature is a major source of the complex temporal patterns that complex adaptive systems generate. (Holland [1995: 10])

Holland's point here is not affected if instead of modeling the agents as simply applying "if/then" rules placed in hierarchical sequences, we model them in more Varelian fashion as acting on more elaborate enacted knowledge about society acquired through interactions with other agents within a society that generates an *esprit critique* which, as the unintended result of these interactions, changes over time. "Adaptation," then, also means something a bit more elaborate than what Resnick (1994), Holland (1995) or Axtell and Epstein (1996) suggest in their respective *StarLogo*, *Echo* and *Sugarscape* models. It means being exposed to (a subset of) all co-existing "complexity-harnessing" options currently present in the society, whether opportunistic, CA-critical, or OS-critical, and modifying one's own complexity-harnessing option as an effect of persuasion, instilled fear, good argument, rhetoric, and so on—so that all contacts of an individual i with another individual j can be understood, *at that level of explanation*, as "irritations" creating a momentary structural coupling between i and j so that i's internal coherence is momentary challenged, his microworld is modified, and his autonomously self-organizing cognitive capacities have to effect a reorganization.

Obviously, the process by which such individuals interact—that is, the succession of vectors of individuals' critical views on society, t_i (with i extending from 1 to n), inducing an aggregate *esprit critique* which, in turn, acts as a systemic constraint on the next vector of individual positions, and so on—cannot be called dialectical in any sense approaching the Marxian materialist theory of history. What might arguably be retained of the term "dialectical" is the manner in which the individuals seek to formulate to themselves the contradictions of their society, the way in which through interacting and individual reflection they seek to integrate what society "is not" into their critical opinions or theories on how society "is"—in short, the way in which they seek to "think dialectically" (Horkheimer [1937a: 181]). *Modeling individuals as dialectical thinkers in no way implies that the emerging social dynamics will obey any "laws of dialectics"*—saying so would involve a patent instance of the fallacy of composition, and this must be avoided at all costs. Dialectical may be a description of individuals' internal cognitive processes (impelled by certain culturally evolved and evolving forms of what Habermas [1965] generically calls "emancipatory interest"), but not of the way they organize their interactions nor of the way their interactions generate the overall dynamics of the system.

If the final or "long-term equilibrium" characteristics of an emancipated society are not to be already imposed on the process of emancipation that leads to them, neither classical-liberal notions of consensus in which everyone agrees at least on a class of emancipatory models (see Rescher [1993]), nor the Habermasian utopia of a public sphere transformed into an arena of "ideally free speech," will do along the dynamic process. The cognitive consonance conducive to consensus as well as the absence of communicative distortions must themselves be part and parcel of what certain individuals—and perhaps, in the long run, an increasing proportion of them—consider to be an emancipated society, and thus these regulative ideals themselves should offer individuals *criteria for the formulation of critical theories* rather than being postulated at the outset. As history shows, most criteria of equity or symmetry such as the principle of the best argument or the notion of impartial evaluation were actually the by-product of historical *struggles* for the attainment of social forms in which authority was no longer blind and contemptuous, and in which "neutrality" was not merely identified with a minority's self-preservation interests (see e.g. Eagleton [1984: 29-43]). True enough, Habermas himself is not claiming that the regulative ideal of undistorted free speech is empirically satisfied at all times—it may in fact never be, or at least have been, fully satisfied—but that, as a *regulative* ideal, it supposedly holds implicit authority even on those who benefit from distorted communication and thus, somewhat mysteriously, serves as a revealing mechanism for the contradiction in their position, almost like a logical law. But even if this is so, the contradiction has to be revealed *by someone acting as an "irritation" to the otherwise self-enclosed cognitive structure of these individuals*. And this can occur only in a process in which all, or at least most, individuals interact directly or indirectly in their respective bouts to present to others what they conceive as the right way to "judge the so-called free world by its own concept of itself, to take a critical attitude towards it and yet to stand by its ideas" (Horkheimer [1968: ix])—in short, a process of interaction between individuals who critically accept the society in which they are interacting.

Adjourning the teleological figures of consensus and undistorted communication, we are left along the way with either the figure of *modus vivendi* or the figure of compromise. The Hobbesian overtones of *modus vivendi* make it unfit for a society in which the emergent property is not a Leviathan but an *esprit critique*—in fact, the latter offers a privileged tool with which to analyze rational compromise as the manner in which individuals endowed with critical instrumental rationality "agree to disagree." This particular agreement is not a relinquishment of

diverging aspirations; rather, it is the mutually agreed-to suspension of overt conflict so as to leave open future arenas for critical reflection. (For a more detailed discussion of this crucial figure of compromise, see Arnsperger and Picavet [2004].) Just as in a market society prices serve as guides for the agents' subsequent decisions, in the society of critical agents I am discussing here the *esprit critique* serves as a guide to the agents' subsequent decisions on how to revise their critical views—it is, to speak in Hayekian terms, a "summary" of the overall critical atmosphere reigning in the society, and a predominantly Marxist *esprit critique* might for instance make it less likely (but not impossible) for an individual to suddenly become a free-marketeer.

A reader of a more Hobbesian persuasion might at this point interject that the word "instrumental" has been underutilized, and that if taken in its full force it will almost certainly lead to a social dynamics in which critically rational agents, seeking to deploy means to reach their ideals, will quickly become deadly enemies as their respective means-ends calculations (oriented though they are by society-scale ideals rather than by mere self-preservation) clash into each other and make frustrations seer. The reply to this objection is direct, however. The same objection could be—and has been, often by misguided critics—raised against Hayek's market model: in a society where individuals seek to maximize their utility or their profit, instrumental calculations will quickly instruct them to eliminate their competitors, or even to kill them if they can get away with it, but at the very least to cheat on them and on their customers, etc. In other words, the market order will quickly self-destruct. The Hayekian answer is immediate: cultural co-evolution of norms, rules and laws which make this behavior suboptimal *in the very interest of instrumental calculation, viewed in the longer term*—the famous "general business climate" of Wall Street newspapers. Since this is an attempt to offer a Hayekian Critical Theory, a similar argument can be constructed for our society of critically instrumental rationalists, and it can be gotten in large part from an idea voiced by Horkheimer: "The idea of self-preservation, the principle which is driving [instrumental], reason to madness [because it causes violence etc.], is the very idea that can save [critical] reason from the same fate. Applied to concrete reality, this means that only a definition of the objective goals of society that includes the purpose of self-preservation of the subject, the respect for individual life, deserves to be called objective" (Horkheimer [1946: 175]). Therefore:

Second Proposition of Hayekian Critical Theory. *A society of critically instrumental rationalists will co-evolve institutions, regulations, rules and norms which ensure that the "agreement to disagree" on how society ought to be and on how to improve it will be sustainable without a self-destruction of the "esprit critique." This implies a self-restriction of all the critical theories and opinions held by the individuals to ensure at least the self-preservation of all individuals and the continued capacity for all individuals to act purposefully on their critical convictions except if this implies that other individuals will no be able to do the same. One of the implications is that no ideal of society can be called "critical" if it is inherently self-realizing (e.g., if part of the means allowed for by the critical theory consist in killing or otherwise incapacitating opponents, etc.).*

This puts on the same footing Mao's version of "Long-March" socialism and Pinochet's vision of market liberalism. Those Hayekians who sneer at this second proposition because it seems to them "obviously unrealistic" should, at the very least, explain why they accept Hayek's doctrine of cultural evolution for a self-guiding social system of individuals endowed with uncritical instrumental rationality, and why they reject a structurally identical doctrine for a self-criticizing social system of individuals endowed with critical instrumental rationality.

One important question still left open is what the exact correspondent is to Hayek's market competition. If prices are the emergent property of a continuous process of (most probably non-Walrasian) competition, what is the *esprit critique* an emergent property of? To put it in nutshell (again, recasting Hayek's own intuitions about the rise of the competitive logic), *it is the emergent property of whatever methods of social coordination a given set of critically instrumental rationalists will co-evolve along the process*. If these individuals progressively interact through competition in all realms—including the realm of critical ideas itself—this will be because the particular "alchemy" between the contents of their critical visions of society and the way these contents have been made to interact (through teaching, writing, speaking, political deliberation and other diffusion mechanisms) have given rise to across-the-board competitive organization as a *conscious, reflective compromise* between these individuals. Thus, saying that all social coordination methods are bound to evolve into competitive mechanisms would be a foregone conclusion—one which would be as teleological as the affirmation of consensus or undistorted communication. (In fact, one can easily see that Habermas's undistorted communication is the *analogon* to a competitive mechanism transferred into the realm of discursive interaction and driven by the criterion of the best argument.) This is perhaps the deepest difference between Hayek's narrow theory of a competitive-market society and what I have called Hayekian Critical Theory:

Third Proposition of Hayekian Critical Theory. *Rules and institutions which in social evolutionism is explained as the unintended emergent outcome of adaptation processes driven by self-preservation become, in Hayekian Critical Theory, the intended compromise outcome of "adaptation" processes driven by critical aspirations. (What remains an unintended emergent outcome of the same processes is the "esprit critique" of the society, which in a free society escapes and should escape the grasp of any individual or group of individuals.)*

3.3. A self-criticizing complex system does not observe itself: In praise of a multiplicity of earnest perspectivists

I must emphasize again that the key emergent property of *esprit critique* is by no means a sort of Hegelian "Absolute-Spirit" entity floating above society and able to embrace it in one single gaze. Rather, the *esprit critique*—perhaps at this point we should offer as a synonym the expression "esprit *de* critique"—of a self-criticizing society is, like market prices in a market society, *a parameter internalized by each individual* in his own interactive quest for a better social world. Each individual *de facto* posits herself as an observer of social reality (through her experience, her inferences from it and from other people's experiences, her readings, etc.), but the way she goes about shaping her observation is systemically constrained by her society's *esprit critique*. Just as in a well-regulated market society you cannot charge a price different from the prevailing market price, in a well-regulated self-criticizing society you cannot interact with other critically-minded fellow individuals in ways that violate the prevailing *esprit critique*.

Now, a pet problem of many systems theorists is, Can a system observe itself? Rephrased in terms of the present framework, the question could be: Isn't what emerges from the various individuals' interacting critical social visions—while being on a different level from the "sum-total" of these visions—the social system's own observation of itself? According to Luhmann, a system's capacity for self-observation is directly connected to its operational closure:

A system can be observed from the standpoint of its environment, if the environment is able to organize and activate such a capacity, or it can be a self-observing system. This decision is central for the sociologist because we have the choice between seeing ourselves as external observers—such as when, for instance, we talk about the economy or about politics and do not intend to thereby make money or do politics—or as internal observers—which we cannot avoid when we do social theory. Indeed, as soon as we seek to communicate, we are already participating in society. Even social criticism, when it analyzing society, has to draw the conclusion that it must always also include itself as one of the ongoing operations. Here (…) we need no third position, no position outside. (…) If we want to understand the self-observation of a system, we have to be able to adopt the viewpoint of the theologian who attempts to convey the belief in God in the mane of God, or the viewpoint of the pedagogue who for obvious reasons has to defend the idea that at the end of the day, education will give good rather than bad results and that we ought to have the courage to trust in it. (…) In this distinction [between self-observation and external observation] it is difficult to see where we should put the subject. This classic figure would in fact not allow that the subject be given a choice being inside the system it is describing or outside of it. The problems begin when we allow for an out-of-this-world, extramundane subject—is the transcendental subject outside of the world? (…) [If it were so, and] if we

ourselves were to partake in this choice, we would have to face a difficulty: the observer has to operate and has to signify something, since it is hard to imagine how one can signify something from outside the world while making use of the distinction between world and non-world (…). (Luhmann [2002]).

The fact that the text verges on a play on words should not detract us from the deep logical problems which, indeed, would be involved if we really attempted to make sense of an observer signifying objects in the world while being literally outside of the world. This does not seem to imply, however, that anyone who attempts to talk about a whole social system while obviously operating within that system is *necessarily* speaking "in the name of" that system, effectively being the spokesman of the system's "self-observation." Luhmann may be correct that this is *sometimes* the case, especially when the system one is speaking about is the eternal substantive totality of all totalities (such as the Spinozian "God") or an abstract, all-encompassing concept (such as "education"). But the crucial aspect of adopting a *critical* viewpoint on society as a whole is that one is establishing a connection between *current society* as an object of critical enacted cognition and *future desired society* as an object of struggle and compromise. Critical instrumental rationality does not claim to be merely "describing" or even merely "observing" society as it is—i.e., society as experienced by anyone who is currently "operating" in it—but to adopt an *anticipatingly descriptive perspective* on society in such a way as to establish a "relation between rational intention and its realization" (Horkheimer [1937b: 217]).

Esprit critique as the sum-total, or rather the emergent property, of such a set of critical perspectives does not at all constitute the social system's "self-observation" because from the interaction between a set of selective, purposefully future-oriented visions an exhaustive bird's-eye view can emerge only by pure chance. There is certainly no higher law of composition involved. Thus, leaving Hegelian schemes aside, what comes out of our discussion is rather the picture of a society of earnest perspectivists: each individual carries in his mind what is *to him* the most adequate way of synthesizing his enacted knowledge about society so as to criticize it, and it is with the firm conviction that his perspective is *a priori* the best that he enters into interaction with the experiences of the other individuals. In other words, while the actual, ongoing social process is a *concrete totality* which effectively (as Hayek would have it) no one can observe and completely embrace in one sweep of the eye, but which nevertheless owes some of its concrete intellectual and institutional features to the purposive critical-mindedness of its members, this purposive critical-mindedness engenders in each individual a *partial totality* which is his critical vision or theory of society. The concrete social process, then, is driven on by the earnest efforts of

the individuals both to convey knowledge to others and to accept knowledge from the others — it is a process which, in some sense, is still "blind" *à la* Hayek, but this "blindness" now coincides with the individuals' sincere attempts to gain a clear an insight into what is wrong with their society and how it could be improved.

But if *that* revised type of "blindness" is really an indispensable feature of any free society, what are the implications for education and, more specifically, for the teaching of social science and economics? How can we move from "the use of pragmatic everyday knowledge in society" — Hayek's narrow model — to "the use of critical knowledge about society in society" — the idea of a Hayekian Critical Theory?

4. On the use of knowledge about society in society: Horizons for theoretical practice and teaching

4.1. *Path-dependence and the exploration-exploitation tradeoff in social science*

Hayek famously stated that social phenomena are "[t]he results of human action but not of human design" (Hayek [1967: 96]) and that "[t]he attitude of the liberal towards society is like that of the gardener who tends a plant and in order to create the conditions most favorable to its growth must know as much as possible about its structure and the way it functions" (Hayek [1944: 18]). By that he has often been taken to mean narrowly that top-down intervention into self-organizing markets can only stifle the spontaneous ordering mechanisms that pivot around competitive price formation and its capacity for socially optimal information dissemination. It is true that he *did* mean that, but our whole preceding discussion leads us to broaden his insight considerably, and not necessarily in the direction his narrow, principled pro-market theory would accept: treating society like a gardener who tends a plant can mean no more than acknowledging that there are *some* emergent phenomena in any free society, and that to the extent a social phenomenon is really emergent — rather than, for instance, a sum-total or a commanded outcome — it simply *cannot*, for epistemic reasons, be "controlled." In other words, knowing as much as possible about the structure of society and the way it functions means no more than knowing where to locate the society's specific emergent phenomena and giving up the illusion of being able to embrace them from the top down in order to steer them voluntaristically. Any society that is free has, so to speak, a set of *characteristic emergents* which characterize the particular way(s) in

which it is a "free" society. That part of Hayek's insights is, as I indicated in the very first sentence of this paper, his definitive and absolute contribution to social science.

This does *not* imply, however, that *all* free societies' characteristic emergents are *bound to be* competitive market prices, or even more broadly competitively generated indices. Reducing the set of admissible characteristic emergents to competitively generated market prices is a step which Hayek, at least in his earlier writings, takes too lightly because by doing so he completely short-circuits the whole complex dynamics outlined in section 3.2 above. Even in his later writings such as *The Constitution of Liberty* (Hayek [1960]), *Law, Legislation and Liberty* (Hayek [1973-79]) or *The Fatal Conceit* (Hayek [1988]), where a broader institutional co-evolutionism is grafted onto the specific rise of markets, one of the constantly underlying axioms of the analysis is that virtually the only "really free" society can be that which co-evolves market institutions and minimally regulative democratic institutions. No one can, of course, question Hayek's argument if it is taken to say that such a co-evolution is *one* possible path; it will also be a desirable one in several respects if, along the dynamic path of critically-minded interaction, most or all agents adopt as their critical theory of society one version or other of libertarian social theory. Indeed, in his specific adherence to free-market ideas Hayek himself believes that every sensible person can only be, upon reflection, a Hayekian—and in that sense his theory of free-market society is indeed a critical theory of society in the sense expounded earlier by Geuss, i.e., a theory which satisfies Geuss's condition (C) "that the transition from the present state to the proposed final state can come about only if the agents adopt the critical theory as their 'self-consciousness' and act on it."

Hayekian Critical Theory as I have presented it here, however, cannot remain content with this selective focus on one critical vision of society. Indeed, a society in which everyone is a Hayekian—or, for that matter, a Marxist or a Rawlsian—is one very special outcome of the process of critically-minded interaction. Thus, in coherence with Hayek's main contribution to social science, what has necessarily precede any focus on a complex *market system* is the focus on a complex *critical-interaction system*. This system can—but isn't simply "bound to"—generate a dynamic path whose steady state, stable or unstable, may be a complex market system. One of the conditions for this is certainly that along the dynamic path a majority of critically-minded agents come to believe that being a Hayekian economic agent is the right way to criticize their current society (whatever that society is at given points along the path); and for

that to be the case, what must happen in turn is that as the actual social arrangement gradually approach the Hayekian market-society ideal, either most of those agents who are still skeptical of it shift from critical acceptance to reflected adhesion, or despite their critical acceptance no longer have any possibility of harnessing the complex interactions with others so as to alter the increasingly Hayekian social arrangements. Thus, as befits a complex system, the dynamics of critically-minded interaction will always be strongly *path-dependent*: an early shift of a critical mass of people to Hayekian doctrine, or to Marxist or Rawlsian doctrine, may create hard-to-reverse tendencies towards the realization of one of the critical models initially in circulation and may make it increasingly difficult for dissenters to do anything else but "play along" while experiencing their critical acceptance as an increasingly sterile inner tension.

Clearly, once one has understood this sort of dynamics—and, to repeat, something like it is simply indispensable in any free society—one realizes that perhaps the key parameter in orienting or harnessing complex critically-minded interactions is to affect the *variety* of circulating critical viewpoints (Axelrod and Cohen [2000: 32-61]) and the ways in which radically opposed viewpoints can be allowed to *interact* (Axelrod and Cohen [2000: 62-116]). Under appropriate conditions, one can hope to progressively generate a society in which most of the important "nodes" in the social network adhere to a given critical theory of society t^*, which will then gradually become self-realizing along the dynamic path of otherwise "free" interactions between critically-minded individuals. This is precisely why, in addition to being a brilliant social theorist, Hayek was also a shrewd political activist. Keith Dixon (1998: 18-39) has retraced his attempts—most notably through the creation of the *Société du Mont-Pèlerin*—to harness social complexity by enticing an elite of key "nodes" into his specific vision of a free-market society. The key difference between this kind of activism and the brutal repression of dissent in, say, Stalinist communism is that the latter used a completely degenerate form of harnessing—namely, the actual breaking-up of interactive complexity by outlawing dissent from a narrow set of "official" doctrines and thus making the emergent *esprit critique* into a non-emergent, decreed *esprit de corps*. In his activism aimed at key actors and politicians, Hayek was trying to harness complexity in an infinitely more subtle way—which does not exonerate his method from critical scrutiny as to the *induced* (rather than decreed) losses of variety and losses in interactive resources which it purported to cause. Viewing society as a complex adaptive system in which critically-minded individuals interact thus allows us to see what are the key differences (in terms

of *ex ante* principles), but also the key similarities (in terms of *ex post* outcomes), between "harnessing of complexity as complexity-destruction" and "harnessing of complexity as shrewd manipulation."

The *path-dependence* inherent in critical-interaction dynamics suggests that an additional dimension of a free society is that, over and above permitting interactions between critically-minded individuals, it should endow itself with institutions designed to ensure that—in formal parallel with the classical economic argument for anti-trust laws—the society's *esprit critique* will emerge from interaction between a *large variety* of critical visions and theories. As Axelrod and Cohen argue, each complex adaptive system possesses its own specific mix between "exploration" and "exploitation": too much exploration distracts agents from exploiting existing acquisitions, but too much exploitation detracts agents from further exploration. Think tanks and ideological agencies of all sides (see e.g. Smith [1989], Grossman and Helpman [2001]) have a systematic bias towards exploitation—of the potentials of their own preferred doctrine—and away from exploration—of alternative doctrines which would jeopardize the domination of their preferred one. One way to describe Hayek's paradoxical existence is that he spent his life arguing in *favor* of industrial market society's explorative potentials—but this led him to militate in high places *against* intellectual exploration of what he conceived of as "wrong" alternative doctrines. The same is true of many other intellectuals, on the Left as well as on the Right.

4.2. *Can critical instrumental rationality be taught? Leaving the "ignorance argument" in its right place*

What has just been said about path-dependence might go a long way in explaining the often unexpected tensions thrown up in academic as well as political circles by methodological discussions. The acuity of the recent French crisis surrounding the teaching of neoclassical economic theory, as well as the acrimony of some of the ensuing debates centering around the controversial label of "post-autistic economics" (see Fullbrook [2003]), seem to the untrained eye to be rather exaggerated, and a matter of a few overblown egos clashing into one another. In fact, as in any complex adaptive system which has settled into an *unstable steady state* and therefore constantly hovers around a menacing *critical threshold*, the promoters of the currently dominating neoclassical teachings see clearly that their numerous ways of harnessing the system's complexity to their doctrinal benefit (through journal networks, prizes, particular self-perpetuating "standards of the profession," etc.; see Axelrod and Cohen [2000: 117-151]) may be

put in jeopardy by an excessive promotion of "criticism" and "alternative schools of thought." In fact, the teaching and/or *de facto* dissemination of non-mainstream theoretical frameworks will, at all times and independently of which framework has currently been brought to the fore by path-dependence, have a tendency to weaken path-dependence and to push the system's acquired steady state (i.e., the dominant framework) on the verge of instability. Tensions constantly build up in the system as one view becomes dominant, and are periodically released as other alternative views acquire a larger share of attention. This can be put as an additional centerpiece of our present analysis:

Fourth Proposition of Hayekian Critical Theory. *The dynamics of a complex adaptive system with critically instrumental rationalists is frequently such that along its path, either one viewpoint on society is rising to dominance through path-dependence (and note that "dominance" does not mean uniqueness or absolute hegemony, just a very large share of "mind-occupation"), or alternative views are gnawing at the dominant view and threatening to destabilize its dominance. The majority of destabilizations will not jeopardize the dominance, but occasionally a large enough destabilization may cause the dominance to vanish, and the complex interactive process to enter a new path-dependent rise to prominence. Thus, the system is usually in a regime of "self-organized criticality" (Bak [1996], Ball [2004: 295-300]).*

It is essential to realize that this self-organized criticality in the "space" of critical views on society makes sense only if the First Proposition above can make sense—if not, then the notion of "self-organization" in that space makes little sense. In particular, the Fourth Proposition obviously presupposes critically rational (rather than instrumentally rational) individuals because in the absence of critical rationality the system's "tension" drops to zero—so that a n ominous corollary follows: if you want a critical theory or a critical view of society, say *S**, to become dominant and then to go unchallenged, you have to succeed in doing at least two things:

(a) As long as *S** is not yet dominant, identify critical reflection with the task of comprehending and mastering *S**.

(b) As soon as *S** has become sufficiently dominant (not everyone agrees on the threshold "mind-occupation" rate that can be deemed sufficient), disparage all critical reflection which does not use the language, axioms and assumptions of *S**.

Such a strategy was certainly part of Hayek's own political agenda. It does not often work perfectly—and Hayek's own failures are an illustration—because interaction-based, embodied and enacted critical reflection has more tricks up its sleeve than the above two-point strategy

assumes. *Critically lived experience*, especially the reflected experience of suffering, of social frustration, and so on, has a knack for destabilizing even the most solid convictions. Yet this strategy is constantly available to well-meaning individuals, or groups of them, who feel their way of critically envisaging society should be promoted to the exclusion of all others.

One straightforward but also notably unstable strategy for countering this hegemonic tendency is *to make critical instrumental rationality itself an object of widespread teaching and practice*. Rather than merely teach one or two dominant views of why and how society should be improved, teach individuals to critically reflect on whatever experience they are living through, to locate the social factors that intervene in that experience, and to shape and reshape their critical view on society by constantly interacting with other similarly-minded individuals. As I explained when arguing in favor of a society of earnest perspectivists, in this interactive probing individuals should not be deterred by what William Brock and David Colander have called the "ignorance argument":

The ignorance argument is the following: the economy is a self-organized system that is beyond our formal modeling capacities; it has emerged through a complex set of interactions. To think that we can actually positively affect something so complex as the market in a positive way is hubris. (Brock and Colander [2000: 82-83])

This argument indeed applies to the concrete totality of the actual, ongoing social process, but it does not apply to the individuals' theoretical models as partial totalities. These may very well—although of course they don't have to, as when an individual uses Epstein and Axtell's *Sugarscape* model as a guide for harnessing his actual society's complexity—be noncomplex or sufficiently simplified to be tractable. So deterring individuals from forming their critical view of society by applying this ignorance argument is not a valid step. However, what is true is that the *esprit critique* that emerges from their interactions, and that acts as a guide for their revisions of their critical viewpoints, *is* subject to an ignorance argument: to think that an individual or set of individuals can affect something so complex as his society's *esprit critique* is hubris—and if an individual or group is to shape the prevailing *esprit critique* by fiat, it will have to be as totalitarian as the individual or group which, in the narrowly Hayekian market society, wants to shape the price vector and/or the distribution of resources by fiat. This is because, in any free society, the *esprit critique* as an emergent phenomenon is part of the concrete social totality, just like the actual ongoing economic process (whether it be a market process or something else).

4.3. Uncompromising ideals, compromising interactions: To which "real world" do we aspire?

Thus, in a society made of interacting, critically instrumental rationalists, the "real world" contains *both* complex social processes as analyzed by contemporary social physics (see Ball [2004]) *and* other complex processes linked to additional dynamic tensions that arise specifically out of the individuals' critical rationality. Obviously, these two sets of processes are separable only in extremely special cases; in most cases, they are not, as evidenced by the fact that—for instance—a Marxist individual and a Hayekian individual may not follow the same rules of behavior *within the same imperfectly competitive market processes that make up their shared concrete social totality.* The Marxist may skew his choices and actions in a given direction due to his lamentation that he is currently living in a *market* society, whereas the Hayekian may skew his own choices and actions in another direction due to his dissatisfaction with living in an *imperfectly* competitive market economy. As two earnest perspectivists, they face the same concrete reality from two irreconcilable perspectives, and their shared "critical acceptance" of that reality makes them coexist (without trying to brutally eliminate each other) while at the same time acting in all ways possible to further their respective emancipatory interests.

This makes the "real world" a complex place on at least two counts: first, its interactive processes generate emergent phenomena due to *instrumental* adaptations in many social arenas; second, one emergent phenomenon anchored in *non-instrumental* adaptation is that individuals "agree to disagree" by striking an ever precarious compromises which allow them coexist while keeping their critical views radically distinct and trying to further their respective views through action within that compromise-ridden society. Thus, they have uncompromising critical ideals but they pursue these ideals in a compromising way, including through attempts to entice others into their own ideals. True enough, this duality may mean the death, or the radical mutation, of some of the most traditional critical ideals—the recent attempts to "recycle" Marxism (Van Parijs [1993]) are, I believe, a symptom of this; yet, the resulting idea of a complex adaptive social system in which one of the crucial things individuals interact about is their critical stance towards prevailing social arrangements is a real world to which I think we can all aspire.

References

Adorno, Th. (1962), "Zur Logik der Sozialwissenschaften", reprinted in *Soziologische Schriften I*, Frankfurt, Suhrkamp, 1972, pp. 547-565.

Albert, M. (2003), *Parecon (Participatory Economics)*, London, Verso.

Arnsperger, Ch. (2004a), "Raison des Lumières et critique sociale. 1° partie: Sur la tâche émancipatrice de la philosophie politique et de la théorie économique", *Cahiers du GRISE,* forthcoming.

Arnsperger, Ch. (2004b), "Raison des Lumières et critique sociale. 2° partie: Vers un modèle de rationalité instrumentale critique", *Cahiers du GRISE,* forthcoming.

Arnsperger, Ch. (2004c), "The Use of Knowledge About Society. Part I — Reopening the Road from Frankfurt to Vienna: Why 'Hayekian Critical Theory' is Not an Oxymoron", Chaire Hoover Document DOCH 139 (www.etes.ucl.ac.be) and GREC Working Paper 2004-04 (www.unites.uqam.ca/philo/pdf/Arnsperger_C_2004-04.pdf).

Arnsperger, Ch., and E. Picavet (2004), "More Than *Modus Vivendi*, Less Than Overlapping Consensus: Towards a Political Theory of Social Compromise", *Social Science Information* **43**(2), 167-204.

Arnsperger, Ch., and Y. Varoufakis (2003), "Toward a Theory of Solidarity", *Erkenntnis* **59**, 155-188.

Atlan, H. (1991), *Oui, non, peut-être:Éducation et vérité*, Paris, Seuil.

Axelrod, R., and M.D. Cohen (2000), *Harnessing Complexity: Organizational Implications of a Scientific Frontier*, New York, Basic Books.

Bak, P. (1996), *How Nature Works: The Science of Self-Organized Criticality*, Berlin, Springer.

Ball, Ph. (2004), *Critical Mass: How One Thing Leads to Another*, London, William Heinemann.

Bohman, J. (1996), "Critical Theory and Democracy", in D.M. Rasmussen (ed.), *The Handbook of Critical Theory*, Oxford, Blackwell,pp. 190-215.

Boudon, R. (2003), *Raison, bonnes raisons*, Paris, Presses Universitaires de France.

Dixon, K. (1998), *Les Évangélistes du marché: Les intellectuels britanniques et le néolibéralisme*, Paris, Raisons d'agir.

Dupuy, J.-P. (1992a), *Introduction aux science sociales: Logique des phénomènes collectifs*, Paris, Ellipses.

Dupuy, J.-P. (1992b), "Friedrich Hayek ou la justice noyée dans la complexité sociale", in *Le Sacrifice et l'envie: Le libéralisme aux prises avec la justice sociale*, Paris, Calmann-Lévy, pp. 241-291.

Eagleton, T. (1984), *The Function of Criticism*, London, Verso.

Epstein, J.M. and R. Axtell (1996), *Growing Artificial Societies: Social Science from the Bottom Up,* Cambridge, MIT Press.

Fukuyama, F. (1992), *The End of History and the Last Man*, New York, Free Press.

Fullbrook, E., ed. (2003), *The Crisis in Economics: The Post-Autistic Economics Movement—The First 600 Days*, London, Routledge.

Geuss, R. (1981), *The Idea of a Critical Theory: Jürgen Habermas and the Frankfurt School*, Cambridge, Cambridge University Press.

Grossman, G.M., and E. Helpman (2001), *Special Interest Politics*, Cambridge, MIT Press.

Habermas, J. (1965), "Erkenntnis und Interesse", reprinted in *Technik und Wissenschaft als "Ideologie"*, Frankfurt, Suhrkamp, pp. 146-168.

Hayek, F.A. (1944), *The Road to Serfdom*, London, Routledge.

Hayek, F.A. (1945a), "Individualism: True and False", reprinted in *Individualism and Economic Order*, Chicago, University of Chicago Press, 1948, pp. 1-32.

Hayek, F.A. (1945b), "The Use of Knowledge in Society", reprinted in *Individualism and Economic Order*, Chicago, University of Chicago Press, 1948, pp. 77-91.

Hayek, F.A. (1952), *The Counter-Revolution of Science: Studies on the Abuse of Reason*, New York, Free Press.

Hayek, F.A. (1967), "The Result of Human Action but not of Human Design", preprinted in *Studies in Philosophy, Economics and Politics*, London, Routledge, 1966, pp. 22-4296-105.

Hayek, F.A. (1973-79), *Law, Legislation and Liberty*, London, Routledge.

Hayek, F.A. (1988), *The Fatal Conceit: The Errors of Socialism*, Chicago, University of Chicago Press.

Horkheimer, M. (1937a), "Traditional and Critical Theory", reprinted in *Critical Theory: Selected Essays*, New York, Continuum, 1972, pp. 188-252.

Horkheimer, M. (1937b), "The Latest Attack on Metaphysics", reprinted in *Critical Theory: Selected Essays*, New York, Continuum, 1972, pp. 132-187.

Horkheimer, M. (1946), *Eclipse of Reason*, New York, Continuum, 1974.

Horkheimer, M. (1968), "Preface" to *Critical Theory: Selected Essays*, New York, Continuum, 1972, pp. v-x.

Hirschman, A.O. (1970), *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*, Cambridge, Harvard University Press.

Holland, J. (1995), *Hidden Order: How Adaptation Builds Complexity*, Cambridge, Perseus.

Kelly, K. (1994), *Out of Control: The New Biology of Machines, Social Systems, and the Economic World*, Cambridge, Perseus.

Lakoff, G., and M. Johnson (1999), *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, New York, Basic Books.

Lindblom, Ch. (1990), *Inquiry and Change: The Troubled Attempt to Understand and Shape Society*, New Haven, Yale University Press.

Lindblom, Ch. (2000), *The Market System: What It Is, How It Works, What to Make of It*, New Haven, Yale University Press.

Lindblom, Ch. and D. Cohen (1979), *Usable Knowledge: Social Science and Social Problem Solving*, New Haven, Yale University Press.

Livet, P. (2003), *Émotions et rationalité morale*, Paris, Presses Universitaires de France.

Luhmann, N. (1975), *Soziale Systeme*, Frankfurt, Suhrkamp.

Luhmann, N. (2002), *Einführung in die Systemtheorie*, Heidelberg, Carl-Auer Verlag.

Maturana, H., and F. Varela (1987), *The Tree of Knowledge: The Biological Roots of Human Understanding*, Boston, Shambhala.

Menke, Ch. (1996), "Critical Theory and Tragic Knowledge", in D.M. Rasmussen (ed.), *The Handbook of Critical Theory*, Oxford, Blackwell, pp. 57-73.

Popper, K. (1962), "The Logic of the Social Sciences", reprinted in *In Search of a Better World: Lectures and Essays from Thirty Years*, London, Routledge, 1994, pp. 64-81.

Resnick, M. (1994), *Turtles, Termites, and Traffic Jams: Explorations in Massively Parallel Microworlds*, Cambridge, MIT Press.

Said, E. (1994), *Representations of the Intellectual : The 1993 Reid Lectures*, New York, Pantheon.

Sen, A. (1977), "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory", *Philosophy and Public Affairs* **6**, 317-344.

Simon, H. (1996), *The Sciences of the Artificial*, Third Edition, Cambridge, MIT Press.

Smith, J.A. (1989), "Think Tanks and the Politics of Ideas", in D.C. Colander and A.W. Coats (eds), *The Spread of Economic Ideas*, Cambridge, Cambridge University Press, pp. 177-194.

Taylor, Ch. (1983), "Social Theory as Practice", reprinted in *Philosophy and the Human Sciences: Philosophical Papers, Vol. 2*, pp. 91-115.

Van Parijs, Ph. (1993), *Marxism Recycled*, Cambridge, Cambridge University Press.

Varela, F. (1983), "L'auto-organisation: De l'apparence au mécanisme", in P. Dumouchel and J.-P. Dupuy (eds), *L'Auto-organisation: De la physique au politique*, Paris, Seuil, pp. 147-166.

Varela, F. (1999), *Ethical Know-How: Action, Wisdom, and Cognition*, Stanford, Stanford University Press.

Varela, F., E. Thompson and E. Rosch (1991), *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MIT Press.

Walzer, M. (1987), *Interpretation and Social Criticism*, Cambridge, Harvard University Press.

Walzer, M. (1988), *The Company of Critics : Social Criticism and Political Commitment in the Twentieth Century*, New York, Basic Books.

NUMÉROS RÉCENTS