



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

C. R. Geoscience 337 (2005) 457–474



<http://france.elsevier.com/direct/CRAS2A/>

External Geophysics, Climate and Environment

Living with uncertainty: from the precautionary principle to the methodology of ongoing normative assessment

Jean-Pierre Dupuy, Alexei Grinbaum *

CREA, École polytechnique, 1, rue Descartes, 75005 Paris, France

Received 16 June 2004; accepted after revision 29 December 2004

Written on invitation of the Editorial Board

Abstract

The analysis of our epistemic situation regarding singular events, such as abrupt climate change, shows essential limitations in the traditional modes of dealing with uncertainty. Typical cognitive barriers lead to the paralysis of action. What is needed is taking seriously the reality of the future. We argue for the application of the methodology of *ongoing normative assessment*. We show that it is, paradoxically, a matter of forming a project on the basis of a fixed future *which one does not want*, and this in a coordinated way at the level of social institutions. *Ongoing* assessment may be viewed as a prescription to *live with* uncertainty, in a particular sense of the term, in order for a future catastrophe not to occur. The assessment is necessarily *normative* in that it must include the anticipation of a retrospective ethical judgment on present choices (notion of moral luck). **To cite this article:** *J.-P. Dupuy, A. Grinbaum, C. R. Geoscience 337 (2005).*

© 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.

Résumé

Vivre avec l'incertitude : du principe de précaution à la méthodologie de l'évaluation normative continue. L'analyse de notre situation épistémique face à des événements uniques, comme un changement climatique brutal, met en évidence des limites essentielles dans nos méthodes traditionnelles de traitement de l'incertitude. Des barrières cognitives spécifiques conduisent à une paralysie de l'action. Ce qui est nécessaire, c'est de prendre au sérieux la réalité de l'avenir. Nous plaçons pour l'application de la méthode de *l'évaluation normative continue*. Nous montrons qu'il s'agit paradoxalement de former un projet sur la base d'un avenir *dont on ne veut pas*, et cela de manière coordonnée au niveau des institutions sociales. L'évaluation *continue* peut être vue comme la prescription de *vivre avec* l'incertitude, ce terme étant pris dans un sens particulier, en vue d'éviter qu'une catastrophe future ne se produise. L'évaluation est nécessairement *normative* en ceci qu'elle fait intervenir l'anticipation d'un jugement éthique rétrospectif sur les choix actuels (notion de fortune morale). **Pour citer cet article :** *J.-P. Dupuy, A. Grinbaum, C. R. Geoscience 337 (2005).*

© 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.

* Corresponding author.

E-mail addresses: jpdupuy@stanford.edu (J.-P. Dupuy), grinbaum@poly.polytechnique.fr (A. Grinbaum).

Keywords: Climatic change; Uncertainties; Cognitive barriers; Ongoing normative assessment

Mots-clés: Changement climatique ; Incertitudes ; Barrières cognitives ; Évaluation normative continue

1. Our epistemic situation

1.1. Knowledge is not enough

As a starting point for our reflection, we take as given the three following points. We are aware that each of them still raises heated controversies. However, it is not our aim to discuss the validity of those three points here, but rather to analyse the situation that results from their conjunction for all parties, their number growing, which are convinced of their truth.

- (A1) Global warming due to the accumulation of greenhouse gases in the atmosphere is a reality; its most proximate cause is human activities; as a consequence, calamitous climate change is a dire prospect that cannot be ruled out if substantial reductions in carbon emissions are not implemented without delay.
- (A2) There remain large uncertainties regarding the future. The modelling of climate change, for instance, cannot tell us whether the temperature of the planet by 2100 will have increased by 1.5 or 6 °C. However, it must be noted that half of that uncertainty results from the uncertainty regarding the type of policy that will be implemented, which may vary from a strong determined action to a lack of measures to reduce greenhouse-gas pollution.
- (A3) Although the previous two points have been known for some time, no serious action up to the challenge is being taken, whether on the part of governments, corporations or ordinary people, as of the day when we are writing this. The Kyoto protocol remains arguably a much weaker measure than required. Knowing about the impending threats is obviously not sufficient to prompting a significant change in behaviour.

Humankind is at a crossroad. It knows enough to be aware that if it does not change its course quickly, its very future is in jeopardy. Nevertheless, it does not seem able to bring itself to act on that knowledge, as if

it were in a state of denial. There is no question that psychosocial factors play an essential role here. On the basis of numerous examples, an English researcher David Fleming identified what he called the “inverse principle of risk evaluation” [3]: the propensity of a community to recognize the existence of a risk seems to be determined by the extent to which it thinks that solutions exist. If the implications in terms of cultural perspective and social behaviour require a jump into the radically unknown, the risk is flatly denied.

However, we believe that the major obstacles to true awareness and action must be sought at a deeper level, which has to do with the kind of temporality that we humans experience before a major catastrophe. Our cognitive capacities operate differently from what they do in ordinary times, and the metaphysical categories on which reasoning rests are altered radically: modal operators such as the possible and the necessary, or the uncertain and the probable, obey different laws. If that is true, there is no serious study of how we can extirpate ourselves from the current impasse that can afford not to tackle those fundamental philosophical issues.

1.2. Uncertainty revisited

We will focus mainly here on the issue of uncertainty. It appears at two different levels. On the one hand, there is the uncertainty that inevitably affects our predictions of any future state of affairs, in particular the consequences of our actions. This problem, which for a long time pertained to the realm of rational prudence, has become a full-fledged ethical problem. Here we follow the lead of German philosopher Hans Jonas, who cogently explains why we need a radically new ethics to rule our relation to the future in the ‘technological age’ [4]. This ‘Ethics of the Future’ (*Ethik für die Zukunft*) – meaning not a future ethics, but an ethics *for* the future, for the sake of the future, i.e. the future must become the major object of our concern – starts from a philosophical aporia. The new situation that we have to confront was brought about when we became able to ‘act into nature’ [1], that is, to tamper with, and trigger, complex natural

phenomena. There was a time when climate and its manifestations in the form of weather were the paradigmatic case of exogenous randomness: for classical economics, for example, any exogenous shock was ascribed, metaphorically, to a change in weather. We know from A1 that at least in part those changes are the result of human action.

Given the magnitude of our power to act, it is an absolute obligation for us to anticipate the consequences of our actions precisely, assess them, and ground our choices on this assessment. Couched in philosophical parlance, this is tantamount to saying that when the stakes are high, none of the normative ethics that are available is up to the challenge. Virtue ethics is manifestly insufficient, since the problems ahead have very little to do with the fact that scientists or engineers are beyond moral reproach or not. Deontological doctrines do not fare much better since they evaluate the rightness of an action in terms of its conformity to a norm or a rule, for example, to the Kantian categorical imperative: we are now well acquainted with the possibility that ‘good’ (e.g., democratic) procedures lead one into an abyss. As for consequentialism – i.e. the set of doctrines that evaluate an action based on its consequences for all agents concerned – it treats uncertainty as does the theory of expected utility, namely by ascribing probabilities to uncertain outcomes. Hans Jonas argues that doing so has become morally irresponsible. The stakes are so high that we must set our eyes on the worst-case scenario and see to it that it never sees the light of day.

However, the very same reasons that make our obligation to anticipate the future compelling make it impossible for us to do so. Unleashing complex processes is a very perilous activity that leads to a practical aporia: ethics demands certain foreknowledge, a condition that science is unable to fulfil. Indeed, one of the very few unassailable ethical principles is that *ought* implies *can*. There is no obligation to do that which one cannot do. However, we do have here an ardent *obligation* that we *cannot* fulfil: anticipating the future. This cannot but violate one of the foundations of ethics.

However, uncertainty is present at another level as well. When it is stated, as in A2, that half of the uncertainty regarding future global warming is due to the indeterminacy of future policies, it is presupposed that our choices are treated as resulting from some kind of

free will, i.e. as exogenous variables. This assumption is highly questionable, for it is tantamount to forgetting a crucial causal link, already present in A3 and which we call motivational: what we will decide to do depends in part on our representation of the future and, in particular, on the uncertainty that surrounds it. One might, for instance, surmise that the large uncertainty described in A2 had a paralysing effect on taking action.

Uncertainty, therefore, appears to be both cause and consequence. We will show that taking this circular relation seriously may enable us to overcome the obstacle expressed in A3. What is needed is a novel approach to the future, neither a set of scenarios nor a forecast.¹ We submit that what we call *ongoing normative assessment* is a step in that direction. In order to introduce this new concept, we first make a long detour into the classic approaches to the problems raised by uncertainty.

2. Types of risks, categories of uncertainty

Many methodologies have been devised to deal with similar problems. In all of them, rational analysis fails to take into account the presence of that which it classifies as irrational, and due to this important failure these methodologies have repeatedly proved unable to address successfully the risks issues they were meant to manage. More importantly, they cannot possibly account for that which must remain for them a scandal, namely that the dissemination of knowledge is powerless to prompt action. Long ago Aristotle’s *phronesis* was dislodged from its prominent place and replaced with the modern tools of probability calculus, decision theory, the theory of expected utility, cost-benefit analysis, and the like. More qualitative methods, such as futures studies, ‘Prospective’,

¹ Coming from an altogether different direction, we meet the conclusion of the Editorial Comment in [5]: “For climate research to be successful it has to move beyond scenarios, but it is not realistic to expect that this movement will lead straight to forecast. This paper is describing a state somewhere in between, and the appropriate terminology for that state needs to be defined” (pp. 415–416). The Editorial Comment then calls for “the creation of a separate identity for describing the future that is more than a scenario but less than a forecast” (p. 416). We believe that ongoing normative assessment provides just that.

and the scenario method were then developed to assist decision-making. More recently, the precautionary principle emerged on the international scene with an ambition to rule those cases in which uncertainty is mainly due to the insufficient state of our scientific and technological knowledge. We believe that none of these tools is appropriate for tackling the situation that we are facing nowadays with regard to the future.

2.1. *Objective versus epistemic uncertainty*

The precautionary principle triumphantly entered the arena of methods to ensure prudence. All the fears of our age seem to have found shelter in the word ‘precaution’. Yet, in fact, the conceptual underpinnings of the notion of precaution are extremely fragile.

Let us recall the definition of the precautionary principle formulated in the French Barnier law on the environment (1995): “The absence of certainties, given the current state of scientific and technological knowledge, must not delay the adoption of effective and proportionate preventive measures aimed at forestalling a risk of grave and irreversible damage to the environment at an economically acceptable cost.” This text is torn between the logic of economic calculation and the awareness that the context of decision-making has radically changed. On one side, the familiar and reassuring notions of effectiveness, commensurability and reasonable cost; on the other, the emphasis on the uncertain state of knowledge and the gravity and irreversibility of damage. It would be all too easy to point out that if uncertainty prevails, no one can say what would be a measure proportionate (by what coefficient?) to a damage that is unknown, and of which one therefore cannot say if it will be grave or irreversible; nor can anyone evaluate what adequate prevention would cost; nor say, supposing that this cost turns out to be ‘unacceptable’, how one should go about choosing between the health of the economy and the prevention of the catastrophe.

One serious deficiency, which hampers the notion of precaution, is that it does not properly gauge the type of uncertainty with which we are confronted at present. The Kourilsky–Viney report on the precautionary principle prepared for the French Prime Minister [8] introduces what initially appears to be an interesting distinction between two types of risks: ‘known’ risks and ‘potential’ risks. It is on this distinc-

tion that the difference between prevention and precaution is said to rest: precaution would be to potential risks what prevention is to known risks. A closer look at the report in question reveals (1) that the expression ‘potential risk’ is poorly chosen, and that what it designates is not a risk waiting to be realized, but a hypothetical risk, one that is only a matter of conjecture; (2) that the distinction between known risks and, call them this way, hypothetical risks corresponds to an old standby of economic thought, the distinction that John Maynard Keynes and Frank Knight independently proposed in 1921 between risk and uncertainty. A risk can in principle be quantified in terms of objective probabilities based on observable frequencies; when such quantification is not possible, one enters the realm of uncertainty.

The problem is that economic thought and decision theory underlying it were destined to abandon the distinction between risk and uncertainty as of the 1950s in the wake of the exploit successfully performed by Leonard Savage with the introduction of the concept of subjective probability and the corresponding philosophy of choice under conditions of uncertainty: Bayesianism. In Savage’s approach, probabilities no longer correspond to any sort of objective regularity present in nature, but simply to the coherent sequence of a given agent’s choices. In philosophical language, every uncertainty is treated as *epistemic* uncertainty, meaning an uncertainty associated with the agent’s state of knowledge. It is easy to see that introduction of subjective probabilities erases Knight’s distinction between uncertainty and risk, between risk and the risk of risk, between precaution and prevention. If a probability is unknown, all that happens is that a probability distribution is assigned to it subjectively. Then further probabilities are calculated following the Bayes rule. No difference remains compared to the case where objective probabilities are available from the outset. Uncertainty owing to lack of knowledge is brought down to the same plane as intrinsic uncertainty due to the random nature of the event under consideration. A risk economist and an insurance theorist do not see and cannot see any essential difference between prevention and precaution and, indeed, reduce the latter to the former. In truth, one observes that applications of the precautionary principle generally boil down to little more than a glorified version of cost-benefit analysis.

Our situation with respect to new threats is different from the above-discussed context. Indeed, it merits a special treatment, and this is the fundamental intuition at the basis of the precautionary principle for which it must be credited. The novelty is that uncertainty is objective and not epistemic; however we are not dealing with a random occurrence either. This is because each of the catastrophes that hover threateningly over our future must be treated as a *singular event*. Neither random, nor uncertain in the epistemic sense, the type of future risk that we are confronting is a monster from the standpoint of classic distinctions. Treating all uncertainties as epistemic, the precautionary principle is ill-equipped to tackle it.

2.2. Uncertainty in complex systems

From the point of view of mathematics of complex systems, ecosystems including, one can distinguish several different sources of uncertainty. Some of them appear in almost any analysis of uncertainties; others are taken into account quite rarely. The most widely cited source of uncertainty is the uncertainty in initial data. Let us treat the climate as a complex physical system, each component of which at the micro level obeys the usual laws of physics. A complex system evolves according to a system of differential equations that represent mathematically the above-mentioned laws of physics. Differential equations determine different possible *trajectories* of the system, i.e. paths that its development in time can take. Each trajectory shows with certainty how the system can move from a known point at the present time to some new position in the future. However, one notices that if the initial data are not known with absolute correctness, which is always the case in practice where infinitely precise measurements are not possible, then there appears a margin of error concerning the initial position of the system. This, in turn, leads to not knowing exactly which will be the destination of the system at a future time. Sometimes it happens that a small error on the initial data entails a small error on the final result, which is a nice type of behaviour that one can deal with by using standard concepts of the measurement theory. But in other instances it occurs that the trajectories that start at two points that are initially very close diverge and lead the system in two totally different directions. Then a small error on the initial

data entails a very large uncertainty regarding the final result. In mathematics such systems are called unstable, and instability with respect to initial conditions is the first source of uncertainty. Instability is a characteristic property of chaotic systems and it is responsible for the well-known type of behaviour called *deterministic chaos*.

Deterministic chaos appears due to a conjunction of properties of the complex system, i.e. divergence of its trajectories, and objective properties of the human observer who cannot know the empirical data with infinite precision. Contrary to this, the second source of uncertainty has to do with the intrinsic character of the complex system. Complex systems can exhibit strikingly different behaviour in terms of trajectories. Imagine trajectories as paths drawn on a plane. In one case all paths can be parallel straight lines; or some paths can go in curves, in circles, in spirals or even stranger geometric figures. This picture on the plane can be viewed as a map of a landscape where each curve denotes a possible road for the system to take. The study of the landscape of trajectories is an important task in mathematics of differential equations. It is only possible to know the exact landscape if we know in advance a rather detailed specification of the complex system. When the system is not fully specified, one is left to guess that it may behave one way or another. In particular, it may be so that the system evolves smoothly, in a stable way, and responds robustly to external influences, but that such robustness is exhibited only up to some point. At a crucial moment, the system leaves the range of smoothly running trajectories and jumps into an area of rapid change, abrupt modification of its parameters, where all the silkiness of its previous life will be quickly forgotten. Such discontinuities in mathematics are called *catastrophes* and places where they happen are called *tipping points*. Appearance of catastrophes is a built-in quality of the system due to its landscape of trajectories. Unlike the deterministic chaos, tipping points are not linked to uncertainties in the initial data. Moreover, in terms of the map, it suffices to modify very slightly a drawing consisting of parallel lines, and the landscape of trajectories will be drastically different. Therefore, if a complex system is not fully specified or fully known, it is impossible to state in advance that its evolution is free from catastrophes.

As long as the system remains far from the threshold of the catastrophe, it may be handled with impunity. In the case of the global ecosystem as any other system far from the threshold, cost-benefit analysis of risks appears useless, or bound to produce a result known in advance, since the landscape of trajectories is predictable and no surprises can be expected. That is why humanity was able to blithely ignore, for centuries, the impact of its mode of development on the environment. But as the critical thresholds grow near, cost-benefit analysis, previously useless, becomes meaningless. At that point, it is imperative not to enter the area of critical change at any cost, if one of course wants to avoid the crisis and sustain the smooth development. We see that for reasons having to do, not with a temporary insufficiency of our knowledge, but with objective, structural properties of ecosystems, economic calculation or cost-benefit analysis is of little help.

Thus far the reasoning could be applied to complex systems of various origins. To specialize to the particular example of climate, it has recently been suggested that climatic change may become an issue not of economic concern, but of national security. The report commissioned by the Pentagon and publicized in the newspapers (*The Observer* of 22 February 2004, *The New York Times* of 26 February 2004) makes use of the first and the second sources of uncertainty. To quote, “An imminent scenario of catastrophic climate change is plausible and would challenge United States national security in ways that should be considered immediately... We don’t know exactly where we are in the process. It [a disaster] could start tomorrow and we would not know for another five years.” Here we are far from the area of cost-benefit analysis. The catastrophic change in question is precisely the change due to the passage of a tipping point, to an abrupt fall of the complex ecosystem into the area of a, mathematically speaking, catastrophe.

2.3. *Not knowing that one does not know*

Why is it so important to resist the reduction of all uncertainties to epistemic uncertainty, as Bayesianism, of which the Precautionary Principle is just a specific case, would have it?

When the precautionary principle states that the “absence of certainties, given the current state of sci-

entific and technical knowledge, must not delay etc.”, it is clear that it places itself from the outset within the framework of epistemic uncertainty. The assumption is that we know we are in a situation of uncertainty. It is an axiom of epistemic logic that if I do not know P, then I know that I do not know P. Yet, as soon as we depart from this framework, we must entertain the possibility that we do not know that we do not know something. In cases where uncertainty is such that it entails that uncertainty itself is uncertain, it is impossible to know whether or not the conditions for application of the precautionary principle have been met. If we apply the principle to itself, it will invalidate itself before our eyes.

Moreover, “given the current state of scientific and technical knowledge” implies that a scientific research effort could overcome the uncertainty in question, whose existence is viewed as purely contingent. It is a safe bet that a ‘precautionary policy’ will inevitably include the edict that research efforts must be pursued – as if the gap between what is known and what needs to be known could be filled by a supplementary effort on the part of the knowing subject. But it is not uncommon to encounter cases in which the progress of knowledge comports an increase in uncertainty for the decision-maker, a thing inconceivable within the framework of epistemic uncertainty. Sometimes, to learn more is to discover hidden complexities that make us realize that the mastery we thought we had over phenomena was in part illusory.

3. Uncertainty in self-referential systems

3.1. *Society is a participant*

A new source of objective uncertainty appears in the case of systems whose development participates in the human society. Global ecosystem is just one paradigmatic example. To these systems, the usual techniques for anticipating the future are inapplicable. The difficulty comes from the fact that, in general, any system where the society plays an active role is characterized by the impossibility to dissociate the physical part of the system from the ‘active observer’, who in turn is influenced by the system and must be viewed as one of its components. In a usual setting, the observer looks at the physical system that he studies from

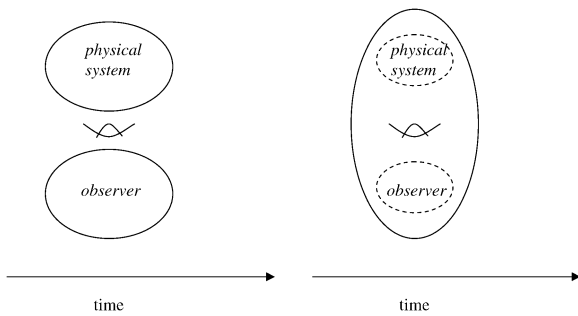


Fig. 1. An external observer and an observer-participant.

Fig. 1. Un observateur extérieur et un observateur participant.

an external point, and both the observer and the physical system evolve in linear physical time. The observer can then treat the system as independent from the act of observation and can create models in which this system will evolve in linear time. Not so if the observer can influence the system and, in turn, be influenced by it (Fig. 1). What evolves as a whole in linear time is now a conglomerate, a composite system consisting of both the physical system and the observer. However, the evolution of the composite system in linear time becomes of no interest for us, for the act of observation is performed by the observer who is a part of the composite system; the observer himself is now *inside* the big whole, and his point of view is no more an external one. The essential difference is that the observer and the physical system enter into a network of complex relations with each other, due to mutual influence. In science such composite systems are referred to as *self-referential systems*. They were first studied by von Neumann in his famous book on the theory of self-reproducing automata, which consequently gave rise to a whole new direction of mathematical research.

Two facts about self-referential systems are to be remembered. The first is a theorem proved by Thomas Breuer, which states that the observer involved in a self-referential system can never have full information on the state of the system. This is a fundamental source of uncertainty in the analysis of complex systems that are influenced by human action. One demonstrates mathematically that, because the human society is capable of influencing the environment in which it lives, it cannot know with certainty the state of the environment nor make any certain predictions about its complete future states.

The other important result is as follows (Fig. 1). In the first case, depicted on the left, the observer builds a sequence of his observations of the state of the physical system in linear time t . He will then be able to construct a database of separate observations. However, in the second case, depicted on the right, due to mutual influences between the physical system and the observer, the observer becomes an *observer-participant* and cannot do his neutral database-collecting work any more. What the observer sees of the physical system reflects not a neutral evolution in linear time, but a highly complex interplay of relations within the composite self-referential system.

One can draw an analogy with changing a frame of reference in kinematics. If you as an observer look at the Earth while standing on its surface, the Earth is at rest, it does not move. In fact, it will never move, because you are in the same reference frame. You can now predict with certainty that at any future time the Earth will stay at rest. However, if you embark on a spacecraft or travel to a space station, the Earth will be moving with respect to your position. Its trajectory will be quite complex: first, when you are going up to space from the surface, it moves back from you, then, when you are on the orbit, it turns around. You need a new special theory if you want to make predictions about the Earth's behaviour. Something quite similar happens in self-referential systems: to put together his observations of the physical system and to put them in accord with his account of mutual influence between himself and the system, the observer needs a new complex theory that depends on the particular character of the observer's role within the *big whole* that it forms together with the physical system.

3.2. Projected time

As a consequence, it is a gross and misleading simplification to treat the climate and the global ecosystem as if they were a physical dynamical system. Human actions influence the climate, and global warming is partly a result of human activity. The decisions that will be made or not, such as the follow-up of the Kyoto protocol, may have a major impact on the evolution of the climate at the planetary level. Depending on whether or not humankind succeeds in mitigating future emissions of greenhouse gases and stabilizing their atmospheric concentration, major catastrophes

will occur or be averted. An organization such as IPCC would have no *raison d'être* otherwise. If many scientists and experts ponder over the determinants of climate change, it is not only out of a love for science and knowledge; rather, it is because they wish to exert an influence on the actions that will be taken by the politicians and, beyond, the peoples themselves. The experts see themselves as capable of changing, if not directly the climate, at least the climate of opinion.

These observations may sound trivial. It is all the more striking that they are not taken into account, most of the time, when it comes to anticipating the evolution of the climate. When they are, it is in the manner of control theory: human decision is treated as a parameter, an independent or exogenous variable, and not as an endogenous variable. Then, as stated at the beginning, a crucial causal link is missing: the motivational link. It is obvious that the human decisions that will be made will depend, at least in part, on the kind of anticipation of the future of the system, this anticipation being made public. And this future will depend, in turn, on the decisions that will be made. A causal loop appears here, that prohibits us from treating human action as an independent variable. Thus, climate and the global ecosystem are systems in which society is a participant.

By and large there are three ways of anticipating the future of a human system, whether purely social or a hybrid of society and the physical world. The first one we call *forecasting*. It treats the system as if it were a purely physical system. This method is legitimate whenever it is obvious that anticipating the future of the system has no effect whatsoever on the future of the system.

The second method we call, in French, *prospec-tive*. Its most common form is the scenario method. Ever since its beginnings the scenario approach has gone to great lengths to distinguish itself from mere forecast or foresight, held to be an extension into the future of trends observed in the past. We can forecast the future state of a physical system, it is said, but not what we shall decide to do. It all started in the 1950s when French philosopher, Gaston Berger, coined the term *prospec-tive* – a substantive formed in analogy with *retrospective* – to designate a new way to relate to the future. That this new way had nothing to do with the project or the ambition of anticipating, that is, *knowing* the future, was clearly expressed in

the following excerpt from a lecture given by another French philosopher, Bertrand de Jouvenel, in 1964. Of *prospec-tive*, he said:

It is unscholarly perforce because there are no facts on the future. Cicero quite rightly contrasted past occurrences and occurrences to come with the contrasted expressions *facta* and *futura*: *facta*, what is accomplished and can be taken as solid; *futura*, what shall come into being, and is as yet 'undone', or fluid. This contrast leads me to assert vigorously: *there can be no science of the future. The future is not the realm of the 'true or false' but the realm of 'possibles'*. [6]

Another term coined by Jouvenel that was promised to a bright future was '*futuribles*', meaning precisely the open diversity of *possible futures*. The exploration of that diversity was to become the scenario approach.

A confusion spoils much of what is being offered as the justification of the scenario approach. On the one hand, the alleged irreducible multiplicity of the *futuribles* is explained as above by the *ontological indeterminacy* of the future: since we 'build', 'invent' the future, there is nothing to know about it. On the other hand, the same multiplicity is interpreted as the inevitable reflection of our inability to know the future *with certainty*. Confusing ontological indeterminacy with epistemic uncertainty in that metaphysical framework is a serious mistake. From what we read in the literature on climate change, we got the clear impression that the emphasis is put on epistemic uncertainty, but only up to the point where human action is introduced: then the scenario method is used to explore the sensitivity of climate change to human action.

The temporality that corresponds to *Prospec-tive* or the scenario approach is the familiar decision tree. We call it *occurring time* (Fig. 2). It embodies the familiar notions that the future is open and the past is fixed. In short, time in this model is the usual linear one-directional time arrow. It immediately comes to mind that, as we have stated above, linear time does not lead to the correct type of observation and prediction if the observer is an *observer-participant*. This is precisely the case with the climate, and, consequently, one must not expect a successful predictive theory of the latter to operate in linear occurring time.

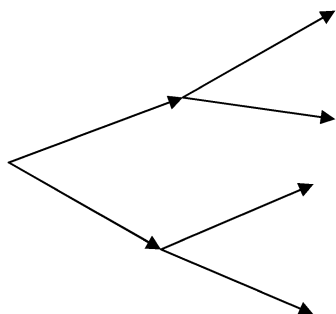


Fig. 2. Occurring time.
Fig. 2. Le temps linéaire.

We submit that occurring time is not the only temporal structure we are familiar with. Another temporal experience – we call it *projected time* – is ours on a daily basis. It is facilitated, encouraged, organised, not to say imposed by numerous features of our social institutions. All around us, more or less authoritative voices are heard that proclaim what the more or less near future will be: the next day’s traffic on the freeway, the result of the upcoming elections, the rates of inflation and growth for the coming year, the changing levels of greenhouse gases, etc. The *futurists* and sundry other prognosticators know full well, as do we, that this future they announce to us as if it were written in the stars is, in fact, a future of our own making. We do not rebel against what could pass for a metaphysical scandal (except, on occasion, in the voting booth). It is the coherence of this mode of coordination with regard to the future that we have endeavoured to bring out.

To return to the three ways of anticipating the future, the foresight method can be said to be a view of an independent observer from outside the physical system. Counter-argument to it is that in reality the observer is not independent and has a capacity to act as to produce causal effects on the system. The second way of anticipation, ‘prospective’, or its version such as the scenario approach, is a view on the system where the observer is not independent any more, but the view itself is still taken from outside the system. Thus, the one who analyses and predicts is the same agent as the one who acts causally on the system. As explained in the previous section, this fact entails a fundamental limit on the capacities of the anticipator. What is needed, therefore, is a replacement of linear occurring

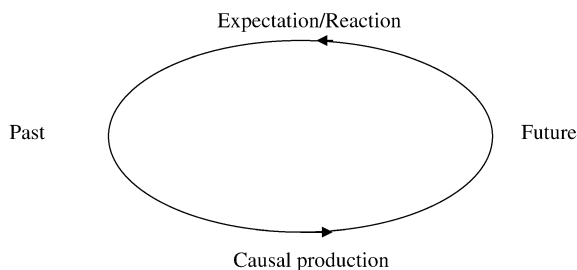


Fig. 3. Projected time.
Fig. 3. Le « temps du projet ».

time with a different point of view. This means taking seriously the fact that the system involves human action and requiring that predictive theory accounts for this. It is only such a theory that will be capable of providing a sound ground for non-self-contradictory, coherent anticipation. A *sine qua non* must be respected for that coherence to be the case: a *closure condition*, as shown in Fig. 3. Projected time takes the form of a loop, in which past and future reciprocally determine each other. It appears that the metaphysics of projected time differs radically from the one that underlies occurring time, as counterfactual relations run counter causal ones: the future is fixed and the past depends counterfactually upon the future.

To foretell the future in projected time, it is necessary to seek the loop’s *fixed point*, where an expectation (on the part of the past with regard to the future) and a causal production (of the future by the past) coincide. The predictor, *knowing that his prediction is going to produce causal effects in the world*, must take account of this fact if he wants the future to confirm what he foretold. Therefore, the point of view of the predictor has more to it than a view of the human agent who merely produces causal effects. By contrast, in the scenario (‘prospective’) approach, the self-realizing prophecy aspect of predictive activity is not taken into account.

We will call *prophecy* the determination of the future in projected time, by reference to the logic of self-fulfilling prophecy. Although the term has religious connotations, let us stress that we are speaking of *prophecy* here in a purely secular and technical sense. The prophet is the one who, prosaically, seeks out the fixed point of the problem, the point where *voluntarism achieves the very thing that fatality dictates*. The prophecy includes itself in its own discourse; it

sees itself realising what it announces as destiny. In this sense, as we said before, prophets are legion in our modern democratic societies, founded on science and technology. What is missing is the realisation that this way of relating to the future, which is neither building, nor inventing or creating it, nor abiding by its necessity, requires a special metaphysics, which is precisely provided by what we call projected time [10].

3.3. *The description of the future causally determines the future*

If the future depends on the way it is anticipated and this anticipation being made public, every determination of the future must take into account the causal consequences of the language that is being used to describe the future and how this language is being received by the general public, how it contributes to shaping public opinion, and how it influences the decision-makers. In other terms, the very description of the future is part and parcel of the causal determinants of the future. This self-referential loop between two distinct levels, the epistemic and the ontological, is the signature of the metaphysics of projected time. Let us observe that this condition provides us in principle with a criterion for determining which kinds of description are acceptable and which are not: the future *under that description* must be a fixed point of the self-referential loop that characterizes projected time.

Any inquiry on the kind of uncertainty proper to the future states of the co-evolution between climate and society must therefore include a study of the linguistic and cognitive channels through which descriptions of the future are made, transmitted, conveyed, received, and made sense of. This is a huge task, and we will limit ourselves here to three dimensions that seem to us of special relevance for the study of climate change: the certainty effect, the aversion to not knowing, and the impossibility to believe. Those cognitive phenomena, although of psychological nature, participate in the objective constitution of uncertainty, because they fully enter into the self-referential determination of the future.

If there is some hope to overcome the obstacles that prevent knowledge from becoming beliefs (convictions) on the basis of which action will be taken, it is, we submit, on this level of the analysis that it can be grounded.

4. Cognitive barriers

4.1. *The certainty effect*

The late Amos Tversky and Nobel Prize winner Daniel Kahneman are famous for having shown that the way a decision problem is described, or ‘framed’, in their terminology, can have a huge influence on the way people solve it. Obviously, this runs counter a basic tenet of the rational choice theory, namely that the preference between options should not reverse when the framing of the decision problem changes. For instance, it is well known that most choices involving gains are risk averse and choices involving losses are risk taking. The same decision problem, whether it is described in terms of gains or in terms of losses, will therefore elicit quite different responses.

The certainty effect and a particularly puzzling variant of it, the pseudo-certainty effect, can best be described with the following experiment carried out by Tversky and Kahneman [14] (numbers in brackets, which in every case sum up to 100%, indicate the fraction of subjects who showed preference for the corresponding option).

Problem 1. Choose the option you prefer:

- (A) a sure win of \$30 [78%]
- (B) 80% chance to win \$45 [22%]

Problem 2. Consider the following two-stage game. In the first stage, there is a 75% chance to end the game without winning anything, and a 25% chance to move into the second stage. If you reach the second stage you have a choice between:

- (C) a sure win of \$30 [74%]
- (D) 80% chance to win \$45 [26%]

Your choice must be made before the game starts. Choose the option you prefer.

Problem 3. Choose the option you prefer:

- (E) 25% chance to win \$30 [42%]
- (F) 20% chance to win \$45 [58%]

A logical analysis shows the three problems to be equivalent. The same probabilities and outcomes characterize Problems 2 and 3. As for 1 and 2, they are obviously the same if step 2 is reached in Problem 2; on the other hand, choosing C or D does not affect the outcome if the game stops at step 1. However, it turns out that if the participants responded in a similar fashion to Problems 1 and 2, their response to Problem 3 was utterly different.

The difference in responses between Problems 1 and 3 illustrates the famous ‘Allais paradox’. In 1952 Maurice Allais, the future French Nobel laureate in economics, demonstrated that people’s preferences (including those of the best rational choice theorists of the time!) systematically violate the axioms of expected utility theory. In particular, departing from certainty by a certain amount – here a reduction in probability of 20%, i.e. from 100 to 80% – has a much larger weight than a similar reduction in the middle of the range – here, a reduction from 25 to 20%. An important lesson to learn, certainty in itself has a special value and desirability, and people are prone to make special efforts and pay a sufficient price to reach certainty. This is why they prefer A to B in Problem 1 (although the expected utility of B is higher), while they prefer F to E in Problem 3. Tversky and Kahneman called this effect the ‘certainty effect’. They observed in passing that “certainty exaggerates the aversiveness of losses that are certain relative to losses that are merely probable”.

The difference in responses between Problems 2 and 3 is even more troubling. It can only be explained by the fact that people treat Problem 2 in the same manner as Problem 1, that is, the certainty effect determines their choice. However, the certainty here is purely illusory, as the gain associated with option C is contingent upon reaching the second stage of the game. Tversky and Kahneman talk here of a “pseudo-certainty effect” and introduce a quasi-oxymoronic phrase “contingent certainty” [7].

Most importantly, Tversky and Kahneman show that “contingently certain outcomes” play a fundamental role in many negotiations, situations of bargaining or trade-off between several conflicting interests or values. Let us consider for instance the political debate within a country that weighs the *pros* and the *cons* associated with the decision to yield occupied territories. In case of war with its neighbours, those

territories represent an invaluable asset that will surely contribute to victory. On the other hand, yielding the territories would reduce the probability of war, but in a measure that remains fundamentally uncertain. It is very likely that the parties in favour of retaining the asset will have the upper hand in the political debate, because of *the superiority of contingent certainty over mere probability* – as was the case of option C compared to option E.

The relevance of this discussion for the debate on climate change is straightforward. There are those who consider that the economy should not be sacrificed for the sake of the environment: in case of an impending major climatic disaster, a strong economy would constitute a sure asset to fight its harmful consequences or even to thwart it. Hence the ‘no-regret’ strategy: never consent to an expenditure in the name of the environment that you might have a chance to regret if it turns out that it was made in vain. On the other hand, others plead that strengthening the economy entails an increase in the probability that a major disaster occurs. Again, it is likely that contingent certainty will have the upper hand over mere probability.

Is it possible to ‘demystify’ the pseudo-certainty effect and expose contingent certainty for what it is, a mere cognitive illusion? Formally, it is a matter, whenever possible, of framing decision problems under formulation 3 rather than 2. But then we stumble upon another obstacle: people are in general very bad at handling probabilities comprised between 0 and 1, as opposed to modalities expressed in language, such as certainty or impossibility.

Another famous experiment carried out by Leda Cosmides and John Tooby of the University of California at Santa Barbara illustrates that point cogently. They asked a group of respondents, including medical doctors, the following question: “If a test to detect a disease, whose prevalence is 1/1000, has a false-positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming you know nothing about the person’s symptoms?” Most respondents, including the doctors, said 95%. The correct answer is 2%, as a straightforward Bayesian analysis shows. However, when put in frequentist terms, the same problem is solved readily by most respondents. Over 1000 people tested, 1 is likely to have the disease and 50 others to have a positive result. Only 1 out of 51 tested positive is sick.

If the human mind is a frequentist device, as claimed by Cosmides and Tooby, it will be of little help in handling *singular* events, which are by their very nature unique, such as a major environmental disaster. As a consequence, it will be very difficult to avoid the traps laid by the pseudo-certainty effect.

4.2. Aversion to not knowing

With his 1952 paradox, Allais intended to show that Savage's axioms are very far from what one observes, in economics, in practical decision-making contexts. Soon an example was proposed, a version of which is known under the name of Ellsberg's paradox [2]. The key idea of Allais and, later on, of Ellsberg is that there exists aversion to not knowing. Not knowing must be understood as the opposite of knowing, a negation of a certain ascribed property, and must be differentiated from ignorance. Ignorance presupposes that something can possibly be known, while here we are concerned with a situation of not knowing and not being able to know, because of the game conditions or because of some real-life factors. Aversion to not knowing can take the form of aversion to uncertainty in situations where uncertainty means epistemic uncertainty according to Frank Knight's distinction between risk and uncertainty. However, as a general principle aversion to not knowing exceeds the conceptual limits of Savage's theory.

The Ellsberg paradox is an example of a situation where agents would 'irrationally' prefer the situation with some information to a situation without any information, although it is 'rational' – according to the prescriptions of Savage's theory – to prefer to turn away from information. Consider two urns, A and B (Fig. 4). It is known that in urn A there are exactly seven white balls and seven black balls. About urn B, it is only said that it contains 14 balls, some white and some black. A ball from each urn is to be drawn at random. Free of charge, a person must choose one of the two urns and then place a bet on the colour of the ball that is drawn. According to Savage's theory of decision-making, urn B should be chosen even though the breakdown of balls is not known. If the person is rational in the Savagean sense, she forms probabilities subjectively, and she then places a bet on the subjectively most likely ball colour. If subjective probabilities are not fifty-fifty, a bet on urn B will be strictly

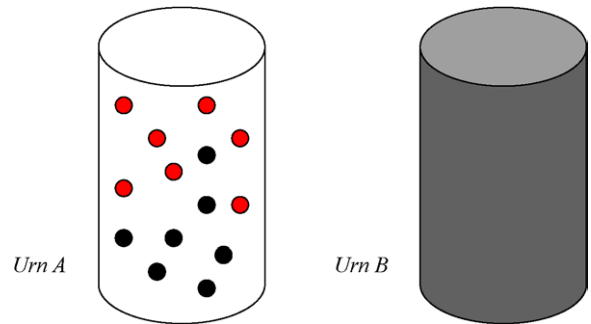


Fig. 4. The Ellsberg paradox.

Fig. 4. Le paradoxe d'Ellsberg.

preferred to one on urn A. If the subjective probabilities are precisely 50:50, then the decision-maker will be indifferent. Contrary to the conclusions of Savage's theory, Ellsberg argued that a strict preference for urn A is plausible because the probability of drawing a white or black ball is known in advance. He surveyed the preferences of an elite group of economists to lend support to this position and found that his view was right and that there was evidence against applicability of Savage's axioms. Thus, the Ellsberg paradox challenges the appropriateness of the theory of subjective probability.

We shall also say that the Ellsberg paradox, along with the Cosmides–Tooby experiment, challenges the usual assumption that human decision-makers are probability calculators. Indeed, had one given himself the task of assessing the problem with urns from the point of view of probabilities, it would be inevitable to make use of the Bayes rule and thus conclude that urn B is the preferred choice. But, as shown by Ellsberg, aversion to not knowing is a stronger force than the tendency to calculate probabilities. Aversion to not knowing therefore erects a cognitive barrier that separates human decision-maker from the field of rational choice theory.

4.3. Impossibility of believing

Let us again return to the precautionary principle. By placing the emphasis on scientific uncertainty, it misconstrues the nature of the obstacle that keeps us from acting in the face of the catastrophe. The obstacle is not just uncertainty, scientific or otherwise; its equally, if not more, important component is the im-

possibility of believing that the worst is going to occur. Contrary to the basic assumption of epistemic logic, one can know that P, but still not believe in P.

Pose the simple question as to what the practice of those who govern us was before the idea of precaution arose. Did they institute policies of *prevention*, the kind of prevention with respect to which precaution is supposed to innovate? Not at all. They simply waited for the catastrophe to occur before taking action – as if its coming into existence constituted the sole factual basis on which it could be legitimately foreseen, too late of course.

Let us consider two examples.

A study of floods in India states:

however, the reason for increasing damages in spite of timely forecasts lies not in forecasting and warning technology, but elsewhere... there is a distinct attitude of indifference towards the warnings [12].

A study on a volcano eruption in Colombia asks:

how did so many people get caught by surprise by Ruiz's catastrophic lahars, in spite of accurate risk assessments and intensive efforts at public education? [13]

These two examples report events that took place on different continents and were linked to climatic predictions of different kinds. Still, in spite of the cultural and natural separation, the effects on the public behaviour are strikingly similar. We submit that there exists a deep cognitive basis for such a behaviour, which is exhibited by human decision-makers in a situation when they know that a singular event, like a catastrophe, stands right behind the door. In these circumstances arises a cognitive barrier of the impossibility to believe in the catastrophe.

To be sure, there are cases where people do see a catastrophe coming and do adjust. That just means that the cognitive barrier in question is not absolute and can be overcome. We will introduce further a method that makes such overcoming more likely. However, by and large, even when it is known that it is going to take place, a catastrophe is not credible.

What could the origin of this cognitive barrier be? Observe first that aversion to not knowing and the im-

possibility to believe do not go unconnected. Both are due to the fact that human action as cognitive decision-making process vitally depends on having information. Cognitive agents cannot act without having information that they rely upon, and the experience from which they build analogies with a current situation. Consequently, a fundamental cognitive barrier arises, which is that if an agent does not have information or experience, then he does not take action, a situation that for an outsider appears as paralysis in decision-making. Aversion to not knowing is caused by the cognitive barrier but the agent, like in the Ellsberg paradox, is forced to act. He then chooses an action that is not rational according to Savage's axioms, but which escapes to the largest degree the situation of not having information. Were the agent allowed not to act at all, as in real life situations, the most probable outcome becomes the one of paralysis. When the choice is between the relatively bad, the unknown, and doing nothing, the last option happens to be the most attractive one. If it is dropped and the choice is just between the relatively bad and the unknown, relatively bad may turn out to be the winner. To summarize, we argue that a consequence of the cognitive barrier is that if, in a situation characterized by absence of information and the singular character of the coming event, there is a possibility not to act, this will be the agent's preference. Having to face the quandary between a catastrophe and a dramatic change in life, most people become paralysed. As cognitive agents, they have no information, no experience, and no practical know-how concerning the singular event, and the cognitive barrier precludes the human decision-maker from acting.

Another consequence of the cognitive barrier is that if an agent is forced to act, then he will do his best to acquire information. Even though it may later be found out that he made wrong decisions or his action was not optimal, in the process of decision-making itself the cognitive barrier dictates that the agent collect as much information as he can get and act upon it. Reluctance to bring in available information or, yet more graphically, refusal to look for information are by themselves special decisions and require that the agent consciously chooses to tackle the problem of the quality and quantity of information that he wants to act upon. If the agent does so, i.e. if he gives himself the task to analyse the problem of necessary vs. superficial information, then it is comprehensible that the agent

would refuse to acquire some information, as does the rational agent in the Ellsberg paradox. But if the meta-analysis of the preconditions of decision-making is not undertaken, then the agent will naturally tend to collect at least some information that is available on the spot. Such is the case in most real life situations. Consequently, the cognitive barrier entails that the directly available information is viewed as relevant to decision-making; if there is no such information, then the first thing to do is to look for one.

Cognitive barrier in its clear-cut form applies to situations where one faces a choice between total absence of information and availability of at least some knowledge. The reason why agents have no information on an event and its consequences is usually that this event is a singular event, like a catastrophe or a climatic change. Singular events, by definition, mean that the agent cannot use his previous experience for analysing the range of possible outcomes and for evaluating particular outcomes in this range. To enter into Savage's rational decision-making process, agents require previous information or experience that allows them to form priors. If information is absent or is such that no previous experiential data is available, the process is easily paralysed. Contrary to the prescription of the theory of subjective probabilities, in a situation of absence of information real cognitive agents do not choose to set priors arbitrarily. To them, selecting probabilities and even starting to think probabilistically without any reason to do so appear as purely irrational and untrustworthy. Independently of the projected positive or negative outcome of a future event, if it is a singular event, then cognitive agents stay away from the realm of subjective probabilistic reasoning and are led to paralysis.

Now, our immediate concern becomes to offer a way of functioning, which is capable of bringing the agents back to operational mode from the dead end of cognitive paralysis.

5. Methodology of ongoing normative assessment

5.1. Reasoning in projected time

Each of the cognitive barriers that we have analysed is an obstacle to reasoning probabilistically, a form of reasoning that presupposes the metaphysics of what

we called 'occurring time'. In that temporality, the future is taken to 'branch out', time takes on the familiar shape of a decision tree, and people reason disjunctively: "I could do this or, alternatively, I could do that"; and "if I were to do this, such consequences would follow"; and "if I were to do that, etc.". We submit that this form of reasoning, in spite of its familiarity:

- is the major obstacle to taking the *reality* of the catastrophe (if there is one ahead) seriously, as it encourages us to turn away from this *reality*, by envisioning other possibilities;
- is not anyhow the way our mind functions spontaneously. Psychologists Eldar Shafir and Amos Tversky conclude their fascinating paper, *Thinking through Uncertainty: Nonconsequential Reasoning and Choice*, as follows:

"A number of factors may contribute to the reluctance to think consequentially. Thinking through an event tree requires people to assume momentarily as true something that may in fact be false. People may be reluctant to make this assumption, especially when another plausible alternative (another branch of the tree) is readily available. It is apparently difficult to devote full attention to each of several branches of an event tree. As a result, people may be reluctant to entertain the various hypothetical branches. [...] The present studies highlight the discrepancy between logical complexity on the one hand and psychological difficulty on the other. In contrast with the 'frame problem', for example, which is trivial for people but exceedingly difficult for Artificial Intelligence [AI], the task of thinking through disjunctions is trivial for AI (which routinely implements 'tree search' and 'path finding' algorithms) but very difficult for people. The failure to reason consequentially may constitute a fundamental difference between natural and artificial intelligence." [11]

It is likely that this 'failure to reason consequentially' may account, at least partly, for the discrepancy between the way experts couch their findings (for instance, in terms of scenarios, another form of an 'event tree') and the reception of their messages by politi-

cians and the general public. We distinguish three paradigmatic types of reasoning:

- (1) non-reflexive reasoning, which may also be called spontaneous decision-making. It does not involve reflection on the rules of reflection, and cognitive barriers rise to their full height. This type is characteristic of the majority of decisions made by human agents;
- (2) mechanistic reasoning of experts and theoreticians, who bring into real life the problematic of ‘event trees’ and scenarios adapted for computer algorithms. Mechanistic reasoning, as shown by Tversky and collaborators, does not have any impact on the spontaneous decision-making of ordinary human beings and does not contribute in the least to the removal of cognitive barriers;
- (3) reasoning in projected time, which we defend here. We submit that this type of reasoning removes the cognitive barriers and can indeed influence non-reflexive reasoning of ordinary people.

Reasoning in projected time has us focus on the future event, taken to be fixed, for example the catastrophe that we need to avert, and avoids diverting our attention on something else. A tension arises, though, that we can formulate as follows: to make the prospect of the catastrophe credible, one must increase the ontological force of its inscription in the future. But to do this with too much success would be to lose sight of the goal, which is precisely to raise awareness and spur action so that the catastrophe *does not take place*.

In the case of a future that one wants to happen, things are simpler. It is then a matter of obtaining through research, public deliberation, and all other means, an image of the future sufficiently optimistic to be desirable and sufficiently credible to trigger the actions that will bring about its own realisation. The tension in this case is between optimism and credibility. However, in the opposite case, the problem becomes one of forming a project on the basis of a fixed future *which one does not want*. If we stated the problem in the following terms: “to obtain through scientific futurology and a meditation on human goals an image of the future sufficiently catastrophic to be repulsive and sufficiently credible to trigger the actions that would block its realization” – such an enterprise would seem to be hobbled from the outset by a prohibitive defect:

self-contradiction. If one succeeds in avoiding the unwanted future, how can one say that a project was formed and action triggered by fixing one’s sight on that same future?

Let us imagine on the other hand that an exceedingly stringent set of policies, like altering radically our way of life, appears to be a necessary condition for avoiding the catastrophe. In occurring time, we have a scenario comprised of hard times in the present followed by a rosy future. It may well be the case that this scenario turns out to be ‘the best’, in that it maximizes expected utility. However, it is unlikely that this sequence of events will constitute the solution to the problem in projected time. Given the democratic setting and the collective psychology of today’s society, the prospect of a rosy future is unlikely to incite people to accept to tighten their belt! What Hans Jonas calls the “heuristics of fear” [5] may well prove a necessary condition for awareness and acceptance of the necessity to change our ways. It is often only when we fear losing something that we value that we become aware of its value for us.

This tension between catastrophism and optimism is inherent in the way the problem of avoiding catastrophes can be solved in projected time. The only way to manage the tension is to imagine that the catastrophe is necessarily inscribed in the future (catastrophism) but with some vanishing, non-zero weight, this being the condition for the catastrophe not to occur (optimism). This means that a human agent is told to *live with* an inscribed catastrophe, and only so will he avoid the occurrence of this catastrophe. Importantly, the vanishing non-zero weight of the catastrophic real future is not the objective probability of the catastrophe and has nothing to do with an assessment of its frequency of occurrence. The catastrophe is altogether inevitable, since it is inscribed in the future: however, if reasoning in projected time is correctly applied, the catastrophe will not occur. A disaster that will not occur must be *lived with* and treated *as if* inevitable: this is the aporia of our human condition in times of impending major threats.

To give an example of how that form of reasoning is applied in actual cases, we cite the Metropolitan Police commissioner Sir John Stevens, who, speaking about future terrorist attacks in London as reflected in his everyday work, said in March 2004: “We do know that

we have actually stopped terrorist attacks happening in London but... there is an *inevitability* that some sort of attack will get through but my job is to make sure that *does not happen*" [9].

5.2. Ongoing normative assessment and moral luck

Projected time is a metaphysical construction, no less so than occurring time. The major obstacle for implementing projected time as the mode of reasoning in our minds and our institutions is that it entails the conflation of past and future, seen as determining each other. However, we are immersed in the flow of time, as the metaphor goes. Linear time is intuitively taken to represent our habitat, and the problem is to project the circular form of reasoning inherent in projected time onto that one-dimensional line that we call time.

This problem justifies why we call for an *ongoing* assessment. The assessment that we are speaking about implies systems where the role of the human observer (individual or collective) is the one of observer-participant. The observer-participant, although it may seem so to him, does not analyze the system that he interacts with in terms of linear time; instead, he is constantly involved in interplay of mutual constraints and interrelations between the system being analyzed and himself. The temporality of this relation is the circular temporality of projected time: if viewed from an external, Archimedes' point, influences go both ways, from the system to the observer and from the observer to the system. The observer, who preserves his identity throughout the whole development and whose point of view is 'from the inside', is bound to reason in a closed-loop temporality, the only one that takes into account the mutual character of the constraints. Now, if one is to transpose the observer's circular vision back into the linearly developing time, one finds that the observer cannot do all his predictive work at one and only one point in time. Circularity of relations within a complex system requires that the observer constantly revise his prediction. To make sure that the loop of interrelations between the system and himself is updated consistently and does not lead to a catastrophic elimination of any major component of either the system in question or of the observer himself, the latter must not stop addressing the question of the future at all times. No fixed-time prediction

conserves its validity due to the circularity and self-referentiality of the complex system.

Reasoning in projected time is meant to ensure that the future is taken to be real. How can this be implemented in linear time? A necessary condition is the coordination of all decision-makers on the same image of the future. Institutions must exist, then, as is already the case for the future of the economy, which give the future of the system under consideration (the climate and its social and cultural underpinnings and consequences in our case) the status of a *focal point*: all actors must take it as a given once they agree on its description. Indeed, some decision-makers, as shows the example of the London police, intuitively apply the methodology of ongoing normative assessment in their everyday work. As long as their number is limited, they all remain individual human agents acting on their own, to whom the term 'coordination' cannot be applied. Collectively coordinating action on a consensual real future, however, is a crucial condition of success of the methodology. What needs to be achieved is a conscious application of the methodology by all parties involved. A reason for this is that the key terms entering in the formulation of the methodology: 'optimistic', 'credible', or 'sufficiently', are initially given a particular meaning separately by each of the parties. If it is not the case that all parties follow the methodology of ongoing normative assessment, thus giving it an institutional status, there will be no common set of concepts in the foundation of their action and, as a result, no possibility of convergence in the understanding of the key terms. Therefore, involvement of all parties in the process of ongoing normative assessment is a crucial criterion of the successful application of the methodology.

A second key characteristic of reasoning in projected time as a way to avoid a future catastrophe is the notion of its inscription in the future with some vanishing, non-zero weight. It gives our paper its title: 'living with uncertainty'. The concept of *that* uncertainty makes sense only in the metaphysics of projected time. What can then be its translation in terms applicable to our linear time, given that projected time freezes, as it were, the familiar flow of time?

It is at this point that our notion of ongoing assessment takes on its specific *normative* dimension. All current approaches to decision-making, including the precautionary principle, share with probabilistic rea-

soning the following feature: the judgment regarding the goodness or the rightness of an action supervenes on the information regarding events that occur up to the moment of that action, and certainly not beyond. In particular, if the consequences of that action were uncertain – which is always the case in practice – only what could be known about them at the time of acting – which excludes their exact determination – will enter into the judgment. In projected time, no such limitation is conceivable, since the time of action entertains no special privilege. It turns out that the concept of ‘moral luck’ in moral philosophy [15] allows us to translate that feature precisely in terms of linear time. We will introduce it by contrasting two thought experiments.

Imagine first that one must reach into an urn containing an indefinite number of balls and pull one out at random. Two thirds of the balls are black and only one third are white. The idea is to bet on the colour of the ball before seeing it. Obviously, one should bet on black. And if one pulls out another ball, one should bet on black again. In fact, one should *always* bet on black, even though one foresees that one out of three times on average this will be an incorrect guess. Suppose that a white ball comes out, so that one discovers that the guess was incorrect. Does this a posteriori discovery justify a retrospective change of mind about the rationality of the bet that one made? No, of course not; one was right to choose black, even if the next ball to come out happened to be white. Where probabilities are concerned, the information as it becomes available can have no conceivable retroactive impact on one’s judgment regarding the rationality of a past decision made in the face of an uncertain or risky future. This is a limitation of probabilistic judgment that has no equivalent in the case of moral judgment.

Examine now the following example devised by the British philosopher Bernard Williams, which we will simplify considerably. A painter – we’ll call him ‘Gauguin’ for the sake of convenience – decides to leave his wife and children and take off for Tahiti in order to live a different life which, he hopes, will allow him to paint the masterpieces that it is his ambition to create. Is he right to do so? Is it moral to do so? Williams defends with great subtlety the thesis that any possible justification of his action can only be retrospective. Only the success or failure of his venture will make it possible for us – and him – to cast judgment. Yet whether

Gauguin becomes a painter of genius or not is in part a matter of luck – the luck of being able to become what one hopes to be. When Gauguin makes his painful decision, he cannot know what, as the saying goes, the future holds in store for him. To say that he is making a bet would be incredibly reductive. With its appearance of paradox, the concept of ‘moral luck’ provides just what was missing in the means at our disposal for describing what is at stake in this type of decision made under conditions of uncertainty.

Like Bernard Williams’ Gauguin, but on an entirely different scale, humanity taken as a collective subject has made a choice in the development of its potential capabilities that brings it under the jurisdiction of moral luck. It may be that its choice will lead to great and irreversible catastrophes; it may be that it will find the means to avert them, to get around them, or to get past them. No one can tell which way it will go. The judgment can only be retrospective. However, it is possible to anticipate, not the judgment itself, but the fact that it must depend on what will be known once the ‘veil of ignorance’ cloaking the future is lifted. Thus, there is still time to insure that our descendants will never be able to say ‘*too late!*’ – a too late that would mean that they find themselves in a situation where no human life worthy of the name is possible. Therefore, moral luck becomes an argument proving that ethics is necessarily a *future ethics*, in Jonas’s sense as described earlier, when it comes to judgment about a future event.

Retrospective character of judgment means that, on the one hand, application of the existing norms for judging facts and, on the other hand, evaluation of new facts for updating the existing norms and creating new ones, are two complementary processes. While the first one is present in almost any sphere of human activity, the second process prevails over the first and acquires an all-important role in the anticipation of the future. What is a norm is being revised continuously, and at the same time this ever-changing normativity is applied to new facts. It is for this reason that the methodology of ongoing assessment requires that the assessment be normative and that the norms themselves be addressed in a continuous way.

From the above discussion, it is clear that no pre-defined norm or rule can be used to ensure the success of the application of the methodology of ongoing normative assessment. Whether it was successful or it

failed, will be judged a posteriori based on the consequences. It is important, therefore, to set up a different criterion of validation, which can only be *procedural* as the methodology itself contains an essential aspect of *ongoingness*. Such a criterion is the agreement of all parties involved on the fact that the prescriptions of the methodology are being applied correctly. It is, as it were, a commonly accepted benchmark of good conduct, where the acceptance itself is being continuously revised. If all parties at all times agree that they correctly follow the procedure, there will arise no sharp tension between them on the questions of conduct. Thus, the methodology of ongoing normative assessment must enjoy a procedural success with all parties concerned; however, whether it will be judged successful at the end, unavoidably depends on the uncertain consequences and is subject to ‘moral luck’.

References

- [1] H. Arendt, *The Human Condition*, University of Chicago Press, Chicago, 1958, pp. 230–232.
- [2] D. Ellsberg, Risk, ambiguity and the Savage axioms, *Q. J. Econ.* 75 (1961) 643–669.
- [3] D. Fleming, The economics of taking care: an evaluation of the precautionary principle, in: D. Freestone, E. Hey (Eds.), *The Precautionary Principle and International Law*, La Haye, Kluwer Law International, 1996.
- [4] H. Jonas, *The Imperative of Responsibility*, in: *Search of an Ethics for the Technological Age*, University of Chicago Press, Chicago, 1985.
- [5] R.N. Jones, Managing uncertainty in climate change projections – Issues for impact assessment, *Clim. Change* 45 (2000) 403–419.
- [6] B. de Jouvenel, *L’art de la conjecture*, Éditions du Rocher, Monaco, 1964.
- [7] D. Kahneman, A. Tversky, Conflict resolution: a cognitive perspective, in: K. Arrow, et al. (Eds.), *Barriers to Conflict Resolution*, Norton, New York, 1993.
- [8] Le Principe de précaution, Report to the French Prime Minister, Éditions Odile Jacob, Paris, 2000.
- [9] Reported by BBC at, http://news.bbc.co.uk/2/hi/uk_news/politics/3515312.stm, Our emphasis.
- [10] (a) J.-P. Dupuy, *Pour un catastrophisme éclairé*, Paris, Seuil, 2002;
 (b) J.-P. Dupuy, Philosophical Foundations of a New Concept of Equilibrium in the Social Sciences: Projected Equilibrium, *Phil. Stud.* 100 (2000) 323–345;
 (c) J.-P. Dupuy, Two temporalities, two rationalities: a new look at Newcomb’s paradox, in: P. Bourguine, B. Walliser (Eds.), *Economics and Cognitive Science*, Pergamon, Oxford, UK, 1992, pp. 191–220;
 (d) J.-P. Dupuy, Common knowledge, common sense, *Theory Decis.* 27 (1989) 37–62;
 (e) J.-P. Dupuy (Ed.), *Self-Deception and Paradoxes of Rationality*, CSLI Publications, Stanford University, CA, USA, 1998.
- [11] E. Shafir, A. Tversky, Thinking through uncertainty: nonconsequential reasoning and choice, *Cogn. Psychol.* 24 (4) (1992) 449–474.
- [12] A. Sharma, Fighting disasters with words! Communication strategies for floods risk reduction in Yamuna river-bed squatters, http://www.gisdevelopment.net/application/natural_hazards/overview/nho0020.htm.
- [13] The Eruption of Nevado Del Ruiz Volcano Colombia, South America, November 13, 1985, Commission on Engineering and Technical Systems, National Academy Press, 1991, p. 58.
- [14] A. Tversky, D. Kahneman, The framing of decisions and the psychology of choice, *Science* 211 (1981) 453–458.
- [15] B. Williams, *Moral Luck*, Cambridge University Press, Cambridge, 1981.