Futures 91 (2017) 62–71

Contents lists available at ScienceDirect

Futures

journal homepage: www.elsevier.com/locate/futures

Original research article

What is wrong with evidence based policy, and how can it be improved?

Andrea Saltelli[a,b,c,*], Mario Giampietro[a,c,d]

[a] *European Centre for Governance in Complexity (ECGC), Barcelona, Spain*
[b] *Centre for the Study of the Sciences and the Humanities (SVT), University of Bergen (UIB), Norway*
[c] *Institut de Ciència i Tecnologia Ambientals (ICTA), Universitat Autonoma de Barcelona (UAB), Spain*
[d] *Institució Catalana de Recerca i Estudis Avançats (ICREA), Spain*

ARTICLE INFO

Article history:
Received 4 September 2015
Received in revised form 5 November 2016
Accepted 7 November 2016
Available online 7 February 2017

Keywords:
Evidence based policy
Science for governance
Science and technology studies
STS
Post-normal science
PNS
Quantitative story telling

ABSTRACT

The present crisis of science's governance, affecting science's reproducibility, scientific peer review and science's integrity, offers a chance to reconsider evidence based policy as it is being practiced at present.

Current evidence based policy exercises entail forms of quantification – often in the form of risk analysis or cost benefit analyses – which aim to optimize one among a set of policy options corresponding to a generally single framing of the issue under consideration. More cogently the deepening of the analysis corresponding to a single view of what the problem is has the effect of distracting from what could be alternative readings. When using evidence based policy those alternative frames become a kind of 'uncomfortable knowledge' which is de facto removed from the policy discourse. All the more so when the analysis is supported by extensive mathematical modelling.

Thus evidence based policy may result in a dramatic simplification of the available perceptions, in flawed policy prescriptions and in the neglect of other relevant world views of legitimate stakeholders. This use of scientific method ultimately generates – rather than resolving – controversies and erodes the institutional trust of the involved actors.

We suggest an alternative approach – which we term quantitative story-telling – which encourages a major effort in the pre-analytic, pre-quantitative phase of the analysis as to map a socially robust universe of possible frames, which represent different lenses through which to perceive what the problem is. This is followed by an analysis where the emphasis in not on confirmatory checks or system optimization but – the opposite – on an attempt to refute the frames if these violate constraints of feasibility (compatibility with processes outside human control); viability (compatibility with processes under human control), and desirability (compatibility with a plurality of normative considerations relevant to the system's actors).

1. Science advice at times of crises

The European Commission is often praised or blamed on reason of being a technocratic organization, where the use of science as input to policy is kept in high consideration. Hence it is not surprising that the symptom of a surging malaise in the

* Corresponding author at: European Centre for Governance in Complexity (ECGC), Barcelona, Spain.
E-mail address: andrea.saltelli@uib.no (A. Saltelli).

http://dx.doi.org/10.1016/j.futures.2016.11.012

use of science for policy are felt acutely in this institution (see Guimarães Pereira and Saltelli, this issue). According to a former EC Chief Science Advisor (Anne Glover, cited in Wilsdon (2014)):

The incoming commission must find better ways of separating evidence-gathering processes from the 'political imperative'.

There seems to be a mounting awareness that too often evidence based policy turns into its opposite, policy based evidence. The recipe to overcome these flaws appears – for those observers – a more strict separation of science and policy.

We intend to argue against this view, and to show that however desirable a pure model of separation between science and policy would appear to be prima facie, it is in practice not viable. The concomitant crises of science, trust and of sustainability call for different medicines that just separating facts from policy. Before we do this though it might be useful to recall what evidence based policy is about.

1.1. What is evidence based policy?

The evidence-based movement began in the field of health, as evidence-based medicine (Pearce, Wesselink, & Colebatch, 2014). Sir Iain Chalmers, a former director of the UK Cochrane Centre, was an early advocate for responsibility in medicine based on a rigorous effort to determine which policies and practices work. Central to this was the concept of experiment, or trial, and in particular the use of randomised controlled trials and the systematic reviews of their results.

While the emphasis of this program was foremost ethical, i.e. to prevent harm from medicine, its main tenet have been criticized (e.g. Hammersley, 2005) on the ground that this approach might lead scientists to make excessive claims for the role. This is not a new critique, in that already in Collingridge and Reeve (1986) warned about the twin myth of rationality that:

1. Policy action can be predicated on the accumulation of facts and the taming of uncertainty; and
2. Science has the power to provide dispassionate facts to adjudicate controversies.

For a discussion of these points see Sarewitz (2000) and Benessia et al. (2016).

Another strand of critique to evidence based policy address its technocratic stance, and its apparent neglect of consideration of power relations: "Policy-relevant facts are the result of an intensive and complex struggle for political and epistemic authority" (Strassheim & Kettunen, 2014). Evidence based policy is indeed used instrumentally to neutralise ideologies and to hide power asymmetries from decision making: "Public conflicts indicate that evidence-based policy rests on its own mechanisms of exclusion and selectivity, that is, modes of black-boxing, knowledge monopolisation, blame avoidance and over-simplification." (*ibidem*). For Boden and Epstein (2006) when using evidence based policy the government (the UK is being considered here) "seeks to capture and control the knowledge producing processes to the point where this type of 'research' might best be described as 'policy-based evidence'.

At present evidence based policy is intended to apply to all policies, not just to health, and to cover all kind of activities, including for example ex-ante impact assessment of policy impact, where science is called to inform policy.

For the specific purpose of the discussion in the present paper we would like to focus our critique in particular to the ex-ante use of evidence for policy. Ex-ante analyses often involve quantification via risk or cost benefit analyses and the use of mathematical modelling to that effect. Such a use of risk and cost benefit analysis – especially when linked to the introduction of new technologies or products – has been at the hearth of the critique of the ecological movement (e.g. Funtowicz & Ravetz, 1994; Schumacher, 1973; Winner, 1986).

Policy based evidence is the flip side of evidence based policy (Strassheim & Kettunen, 2014), and it is impossible to extricate the two, exactly as it is impossible to extricates facts from value when operating at the interface between science (statistics) and policy.

1.2. Controversies

Several cases where science is called to adjudicate a policy seem today to be associated with a growing level of conflict. The provision of scientific input does not seem to quell controversies. For Sarewitz (2000) science seems at time to be the problem rather than the solution: '*Rather than resolving political debate, science often becomes ammunition in partisan squabbling, mobilized selectively by contending sides to bolster their positions.*'

Recent examples of controversies upon which science has been called to adjudicate could include:

- the impact of pesticides on bees,
- the culling of badgers,
- greenhouse potential of the refrigerant liquid used by Mercedes Benz,
- impact of endocrine disruptors,
- benefits of shale gas fracking,
- fate of children raised by gay parents,
- true long term cost of citizenship for illegal migrants,
- desirability of international testing and comparison of the educational attainment of children,

. . . and the list, based on issues that have received mediatic attention in Europe and in the US, could go on.

The point could be made that the term 'wicked issues' (Rittel & Webber, 1973), once reserved to intractable cases such as genetically modified organisms or climate, now applies to most issues where science is called to adjudicate. Climate is so much a defining issue that for Kahan (2015) – studying the impact of group values on the perceptions of risk and of the related evidence – one can determine a person cultural and normative stance by its perception of climate. If this is the case then we must draw the conclusion that our culture presently shape our attitude about a much larger class of problems, and that – as noted by Kahan – the more literate the observer is on a given issue, the higher his polarization is likely to be. It is as if 'facts' – rather than settle the issue – were metabolized instead as ammunition to feed one's vision of the world.

1.3. A crisis in science

Though apparently unrelated, controversy is not helped by science own crisis of reproducibility, integrity and legitimacy. We do not want to offer here a full picture of this crisis, for which that reader is referred to a recent book specifically devoted to this topic (Benessia et al., 2016; the authors of the present paper also contributed to this work). The volume makes the case that the crisis has ethical, epistemological, methodological and even metaphysical dimensions, and that its root causes can be read from history and philosophy of science scholarship to present-day historical critique of commodified science. It is also argued in the book that the crisis of science *qua science* impacts science as used for policy, and that is shown acutely through frictions affecting:

- the paradigm of evidence-based policy;
- the use of science to produce implausibly precise numbers and reassuring techno-scientific imaginaries;
- the use of science to 'compel' decision by the sheer strength of 'facts'.

As per the crisis of science proper – which affect both reproducibility, the peer review system, the use of metrics, the system of incentive, the reader can find a recent summary in the work of Begley and Ioannidis (2015), which summarized a good deal of recent literature including from the same Begley (2013),(e.g. 2013; with Lee, 2012) and Ioannidis (e.g. 2005; 2014). For those authors the ingredients of the crisis are:

- The generation of new data/publications at an unprecedented rate;
- The compelling evidence that the majority of these discoveries will not stand the test of time;
- A failure to adhere to good scientific practice and the desperation to publish or perish and most evident proximate causes;
- The fact that the problem is multifaceted, and involving a large pool of different stakeholders, and that no single party is solely responsible, and no single solution will suffice.

A debate is also unfolding about recent and less recent cases of failure of the science to inform policies, suffice here mention the case of the cholesterol versus sugar saga, where important institutions stood behind misinformed policies for decades in spite of mounting contrary evidence (Teicholz, 2015; see Leslie, 2016, for a summary)

Before all that voices are raised to question whether "science has taken a turn towards darkness" (Richard Horton, editor-in-chief of The Lancet, 2015), or that "Science is turning back to the dark ages" (Phillips, 2016) or what sense of responsibility one might expect from scientists themselves (Macilwain, 2016).

1.4. The roots of the crisis

The present crisis of science did not manifest overnight. At root one of science's crisis could indeed be science's own success. In his 1963 work 'Little Science, Big Science' Derek J. de Solla Price prophesized that Science would reach saturation (and in the worst case senility) under its own weight, victim of its own exponential growth (pp 1–32) (de Solla Price, 1963). de Solla Price is considered today the father of scientometrics. Yet his forewarnings on the impossibility for science to grow for ever, the implicit dangers in the shift from little to big science received relatively less attention. Attention and quality decay due to proliferation of publications in science are today an accepted reality (Della Briotta Parolo, Kumar Pan, Ghosh, Huberman, Kaski, & Fortunato, 2015; Siebert, Machesky, & Insall, 2015), all the while at the same time both so called pure and applied science continues to amaze and inspire with their conquests, from the Rosetta mission landing a probe on a comet flying past the sun to the confirmation of the existence of the Higgs Boson and of the gravitation waves, with artificial intelligences beating champions of Jeopardy first and Go later, and with CRISP-R technologies promising – or threatening? – wonders in the field of genetic manipulation.

Still the crisis is there, and different actors see different relevant or potential causes (Benessia et al., 2016):

- Science victim of an excessive demand from society, poor training, inappropriate statistical design, hubris of data mining, perverse incentives, counterproductive metrics;
- Science victim of its own success, exponential growth determining senility – poor choice of narratives and frames due to lock-in on what worked well in the past – and hyper-specialization;
- Science as another victim of the neoliberal ideology and of commoditization of research changing;

- Science as a social enterprise whose quality control apparatus suffers under the mutated conditions of technoscience.

We cannot tackle all of these in the present short paper, but we need to spend a few words on science as a social enterprise for the deep connection that this reading of the crisis has on the use of science for policy, and for our belief that is aspect is dangerously neglected.

In his 1971 book 'Scientific Knowledge and its Social Problems' Jerome R. Ravetz notes (p. 22):

[. . .] with the industrialization of science, certain changes have occurred which weaken the operation of the traditional mechanism of quality control and direction at the highest level. [. . .]The problem of quality control in science is thus at the centre of the social problems of the industrialized science of the present period. If it fails to resolve this problem [. . .] then the immediate consequences for morale and recruitment will be serious; and those for the survival of science itself, grave

Ravetz identified in the system of quality control the fault line which would run the greatest risk when science – which is a social activity – would find itself with a changed ethos, and when its actors would experience profound changes in the system of reward and incentives. The centrality of ethics for the quality and the self-governance of science so clearly illustrated by Ravetz is also central to the work of Lyotard (1979 Ch. 10). In his 1979 work 'La Condition postmoderne. Rapport sur le savoir' he tackles the de-legitimization of knowledge (identified with science) when this becomes an industrialized commodity – as opposed to the instrument of emancipation and betterment of human beings (*bildung*).

Coming to more recent works, Philip Mirowski describes the degeneration of industrialized science in the US with painstaking detail in his work (Mirowski, 2011) 'Science-Mart: Privatizing American Science'. According to Mirowski after the eighties neoliberal ideologies succeeded in decreasing state intervention in the funding of science, which became increasingly privatized and sub contracted, generating the perverse system of incentive already mentioned by many scholars (e.g. Ioannidis, 2014).

1.5. Science, facts and hybrids

In conclusion of this discussion of science's own crisis we note that the use of science for policy is based on trust, and cannot be independent from what is happening in the science's own house. More specifically the same Lyotard (work cited) notes that since the age of Plato the legitimacy of science is linked to the legitimacy of the legislator:

'*Who decides what counts as knowledge and who knows about what one must decide?* [. . .] *The question of knowledge in the information society is more than ever the question of government*'.

'*Solutions to the problem of knowledge are solutions to the problem of social order*' is a conclusion of Steven Shapin & Simon Schaffer's book 'Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life' (Shapin & Schaffer, 2011). Thus our chosen epistemologies are not just issues for debate in philosophical circles, but have a direct bearing on the polity. This brings us back to Anne Glover's diagnosis and therapy of what needs to be done to fight the sins of 'policy based evidence'. She advocates strengthening the separation between science and policy. This epistemology, known as 'demarcation model' (Funtowicz, 2006), is at present the most widespread, and aims at protecting science from the political interference, preventing possible abuse of science and scientific information driven by agendas. It prescribes a clear demarcation between the institutions (and individuals) who provide the science and those where it is used. Why should we abandon this model?

Our answer to this question is that this well intended model fails the test of realism. The prevailing scientist-policy maker dichotomy, whereby scientists claim to (or are expected to) produce and certify the facts and policy makers claim to guarantee the legitimacy of the values neglects the fact that in present arrangements science's input to policy take place in hybrid settings – with a rich spectrum of actors and competences, from the practicing scientists to the entrepreneur researcher, from the technology regulator to the technical personnel in ministries and international agencies charged with science-laden policy files, often active in boundary organizations operating at the science, law and policy interfaces, under close and eager mediatic scrutiny and interest groups pressure. In these settings, facts and values are intermingled. Instead of purified facts we mostly deals with hybrid arrangements (Latour, 1991), and one of the features of the present epistemic governance crisis is that "*the more knowledge is produced in hybrid arrangements, the more the protagonists will insist on the integrity, even veracity of their findings*" (Grundmann, 2009).

1.6. A word on mathematical models and statistical indicators

Our discussion of science advice at times of crises concludes with an analysis of statistical or mathematical modelling, a fundamental ingredient of science's input to policy, and of the related issue of quantification. We shall argue that in this field the present arrangements are particularly problematic.

We start with a well-known example where science was recruited to advocate austerity in public budgets. A 90% ratio of public debt to gross domestic product was stipulated by Harvard professors Kenneth Rogoff and Carmen Reinhart as an absolute ceiling above which growth would be hampered. Thus debt ratios above this limit were defined as unsafe for a country. A later reanalysis by researchers from the University of Massachusetts at Amherst disproved this finding by tracing it to a coding error in the authors' original work. Clearly once this particular result was repudiated the policies had already be put in place and "*In Britain and Europe, great damage has been done as a result.*" (Cassidy, 2013).

This is but one of the many instances where improper use of mathematical modelling has been instrumental of supporting flawed policies. Modelling hubris and its consequences are discussed in (Saltelli, Guimarães Pereira, van der Sluijs, & Funtowicz, 2013; Saltelli & Funtowicz, 2014). In his 2013 work 'Never Let a Serious Crisis Go to Waste: How Neoliberalism Survived the Financial Meltdown' Philip Mirowski devotes a long section (pp 275–286) to the story of how dynamic stochastic general models (DSGE) were the subject of a hearing in the US senate – '*an event in 2010 that was literally unprecedented in the history of economic thought in America*', p. 275, with sworn testimony of economists such as Sidney Winter, Scott Page, Robert Solow, David Colander and V.V. Chari, to understand how 'theorists tools' had come to be used as policy instruments and why these instruments were all but useless in anticipating the economic crisis. Queen Elisabeth had a similar moment with British economists at the London School of Economics (Pierce, 2008).

Saltelli and Funtowicz (2014) list several problems in the way mathematical modelling is used to tame uncertainty in relation to the production of evidence for policy. These include the rhetorical or ritual use of possibly disproportionate mathematical models to impress or obfuscate, the reliance on tacit possibly unverified assumptions, the instrumental inflation or deflation of uncertainties according to expedience, the instrumental compression and linearization of the analysis as to tame complexity and to convey an impression of prediction and control, and finally an absent or perfunctory sensitivity analysis.

2. From the Cartesian dream to socially constructed ignorance

2.1. From the Cartesian dream to agnotology

Before we move to suggest how the present predicaments of science advice should be tackled we would like to ascertain what needs to be unlearned of our present wisdom in order to achieve progress. We call this the Cartesian dream (Guimarães Pereira & Funtowicz, 2015), and in Benessia et al. (2016) we try to identify elements of the dream from which modernity should perhaps awaken from, starting from Francis Bacon "Knowledge and power meet in one" to recent times dream of science as the Endless Frontier metaphor in Vanevar Bush's vision (1945). Toulmin (1990) calls this dream 'The Hidden Agenda of Modernity', whereby society is modelled on Newtonian physics and hence mastered and controlled by human rationality. The agenda now struggles against the complexities of the present crisis, endangering the legitimacy of the existing social contracts. The fact that the present 'evidence based policy' model clearly subscribes to the Cartesian dream has worrying consequences. The most serious is perhaps in that it induces dramatic simplification. Other terms we could use to describe these simplifications are 'Hypocognition' (Lakoff, 2010), or 'Socially constructed ignorance' (Ravetz, 1986; Rayner, 2012). The study of ignorance or 'agnotology' is also the subject of a successful book by Proctor and Schiebinger (2008).

How does the simplification of evidence based policy manifest itself? This occurs through the mechanism of quantification, which is predicated on a selection of a problem structuring (the adoption of a frame). The selection determines a compression of the aspects that can be considered relevant when observing the external world, and when selecting a limited subset represented in a finite information space. This compression then determines the fragility of the inference based on the chosen simplified representation. The consequence of this process is explained by Rayner (2012) in terms of socially constructed ignorance, which is not the result of a conspiracy but of the sense-making process of individuals and institutions, to produce self-consistent versions of that world by expunging that knowledge which is in tension or contradiction with those versions.

The unavoidable compression associated with the forced choice of a finite representation of a state of affairs comes to a cost, such as ignoring knowledge which is available and is not considered in the given problem structuring ("unknown knowns"), e.g. that knowledge which exists out there in academia but is ignored because in conflict with the goals of institutions or societies.

Other 'victims' of the compression are the "known unknowns" – i.e. one choses to ignore the fact that some elements which would be needed for the analysis are missing.

Once the analysis has removed all sources of uncomfortable knowledge the problem reduces to one which can be treated by the usual combination of cost benefit analysis and risk analyses methodologies, and the solution optimized to the desired precision, be it that the solution may have lost at this stage all its relevance to the original problem.

A 'sustainability' implication of this approach is that the simplification operated in the space of the possible options may have decreased the adaptability of the system, – possible viable options have been removed from the analysis – exasperating the efficiency – resiliency trade off (see Nassim N. Taleb's work 'Antifragile', 2012).

2.2. An example: biofuels

An illustration of these shortcomings, seen from the lenses of bioeconomics, is represented by biofuels. In applications of bioeconomics (Giampietro, Mayumi, & Sorman, 2012; Giampietro, Aspinall, Ramos-Martin, & Bukkens, 2014) it is recommended that simultaneous non-equivalent narratives (dimensions of analysis) and different scales (descriptive domains) are adopted when considering an issue. In the case of bioethanol from corn, neglecting this approach in favour of a reductionist view of the system (based on simplistic energy analysis or dubious economic assessments) resulted in hundreds of billions of tax payer money being spent on an alternative energy source that consumes more or less the same amount of energy carriers that it produces (Giampietro & Mayumi, 2009). This is aggravated by the fact that this strategy increases rather than decreases emission due to changed land use (Fargione, Hill, Tilman, Polasky, & Hawthorne, 2008), and that the narrative of biofuels as a strategy to mitigate emissions neglected that this was achieved, in the models, because of cuts in food production, due to the displacing effect of biofuels versus traditional crops (Searchinger, Edwards, Mulligan, Heimlich, & Plevin, 2015).

What is relevant in this example is the glaring neglecting of available knowledge. In his masterpiece "Bioenergetics and Growth" Brody (1945) already dismissed the option of ethanol as follows: "*It is said that we shall use ethanol and vegetable oils after the petroleum energy has been exhausted. This reminds one of Marie Antoinette's advice to the Paris poor to eat cake when they had no bread. Alcohol and vegetable oils are, of course, more expensive than petroleum fuels. Moreover, if all our crops were converted to alcohol and oil, they would not supply the present rate of petroleum energy consumption*". (Brody, 1945, pag. 968). Without the pressure of powerful lobbies, at that time it was easy for Brody to see the elephant in the room. The spike of economic development associated with the industrial revolution on this planet was possible because of the massive use of oil to save both land – the yields of grain moved from around 1 t/ha to more than 8 t/ha – and labor – the labor productivity of grains of farmers moved from around 1 kg/h to 700 kg/h. The narrative about the advantages of biofuels – assuming a massive use of land and labor to save oil – seems to have missed the mechanism that generated economic growth on this planet in the last two hundred years.

More examples of this predicament in the field of economics are offered in Benessia et al. (2016).

2.3. Mathematical modelling versus plain English

Often the falsification of a frame involving dense mathematical modelling can do without the language of mathematics, but use just plain English. To make an example, a critique of the already mentioned dynamic stochastic general equilibrium models DSGE used as policy instruments (and not as theoretician tool) is possible by falsifying the underlying hypotheses of 'efficient markets' and 'representative agent' (Mirowski, 2013; pp 275–286). This is not a new approach. Translating into English the result of mathematical elaboration was a teaching of Alfred Marshall (Pigou, 1925, p. 427), and although his teaching is echoed by present day economists (Krugman, 2009, p. 9), the general tendency of present day economics is bent toward heavy use of mathematics and deductivism, economics being itself the victim of what Mirowski (1991) calls 'physics envy'. For Reinert (2000), a historian of economic thought, economics has gone thought a full circle: "from scholasticism through innovation and back into mathematical scholasticism" and though "mathematics in economics was indeed a most useful innovation", notably since its introduction in economics in the XVIII century, mathematics is also "one language in which a science may decay to sterility and irrelevance". The use mathematical modelling as 'Latin', to establish authority rather than to inform, is also discussed by Saltelli et al. (2013) in the context of quality appraisal for mathematical modelling.

It is important here to say that the object of the present critique is not modelling as such, but degenerate modelling. The context of use is here important. One thing is to use a model to simulate a policy, another story is the leap whereby the same model is used to justify one (Saltelli & Funtowicz, 2014). The different between the two stages is that the former is achieved by a set of 'caeteris paribus' assumptions which are legitimate in a theoretical pursuit but become reckless in the context of real life policies.

Stiglitz (2011) praises the use of models as *'blinders'* which *'leaving out certain things', [. . .] provide a frame through which we see the world*'. For Krugman (2000) one of the stated functions of economic models is that to discipline modellers' reasoning to accept even counter intuitive results (Krugman, 2000).

A more subtle form of degeneration associated the poor use of models that of 'displacement'. Displacement is for Steve Rayner one of the strategies to socially construct ignorance, and occurs where a model become the end instead of the tool, e.g. an institution choses to monitor and manage the outcome of a model rather than what happens in reality.

We witness today the deployment of mathematical modelling – something referring to an information space semantically closed – to predict the behaviour of reflexive, auto-poietic agents continuously adding new meanings and new relevant aspects to be considered in the description of themselves. The adaptability of reflexive autopoietic [capable and in need of regenerating themselves] systems depends on their ability to operate an information space which is semantically open and continuously expanding (Pattee, 1995). Modelling processes are defended on the basis that they provide rigour, however a careful appreciation of the core as well as auxiliary assumptions needed to run the models would be sufficient to appreciate their irrelevance to real life situations.

We thus look with scepticism to the thesis that modelling approaches which have failed to predict a financial and economic crisis will be able to inform us about the behaviour of a system involving institutions, societies, economies and ecologies, as when applying the DSGE to climate change (Stern, 2016); see our comment in Saltelli et al. (2016).

Again in the field of climate studies the future is with us already when we see models deployed to quantify the impact of increased crime rates (resulting from hotter climate) on the economy (Rhodium Group, 2014; see a criticism in Saltelli, Stark, Becker, and Stano (2015)).

In the opinion of the authors of the present paper describing future climate states and the impact of these on the economy and society – to the point to be expressed in digits – has all the characteristics of what has been called by Weinberg (1972) 'Trans-Science', i.e. something which can be described using the language of science but which science cannot resolve.

This use of quantification will facilitate abuse and corruption and – as noted by Porter (1995) will often be driven by a need for legitimacy by institutions in need of one. Porter also notes the symbiotic relationship between quantification and trust. A result of this is that a climate of trust favours sensible quantifications, while a climate of regulatory confusion or controversy breeds 'mandated' quantification which are less so.

3. Solutions: responsible use of quantitative information

The preceding section was focused on what should be unlearned in order to achieve progress in the use of science in support to policy. What should be learned instead? What are our proposed strategies?

We suggest fighting hypocognition by resisting the fatal attraction of quantification till a robust and heterogeneous set of potential relevant frames has been properly charted. Confronted with a pre-existing (set of) quantification(s) our procedure will be to study how the frames were constructed, and how this pre-analytic choice has then led to a selection of relevant data, indicators and models. We would call our approach Quantitative Story Telling for governance.

3.1. Quantitative story telling

Quantitative story telling (QST) involves a participative and deliberative analysis of the quality of proposed or available policies and narratives on governance.

Note that the expression quantitative story telling lends itself to what analytical philosophers call process/product ambiguity, as one might be referring to a given instance of a story or to the process of telling one; the context should make it clear which is what.

QST proceeds at first and foremost '*via negativa*', using a method of falsification of the available options with respect to:

- feasibility (compatibility with external constraints),
- viability (compatibility with internal constraints) and
- desirability (compatibility with normative values adopted in the given society).

This analysis – to be performed participatively with parsimonious use of mathematical modelling and quantification – will test whether any 'impossibility' or bottleneck can be identified which allows a framing or option to be falsified (in the sense of proven false).

This is not an altogether new approach, and can be associated to Rayner (2012) idea of 'clumsy solutions', meant to accommodate "*unshared epistemological or ethical principles* [. . . in a way that is] *satisficing* [. . .] *rather than optimizing*", and the '*working deliberatively within imperfections*' (van der Sluijs, Petersen, Janssen, Risbey, & Ravetz, 2008) of the Post Normal Science's extended participation model, as well with the 'rediscovery of ignorance' advocated by Ravetz (2015, p. xviii).

A key step in the identification of the feasibility, viability and desirability domains entails looking through different lens – i.e. dimensions and scales of analysis. This strategy for sustainability analysis is detailed in (Giampietro, Mayumi, & Sorman, 2013; Giampietro et al., 2014).

An example of this strategy can be made on food security:

- When checking the feasibility of food security against external constraints (context/black-box – agriculture) we have to measure requirements and supply in terms of kg of potatoes, vegetables and animal products.
- However, if we want to check the viability of food security in relation to internal constraints (black box/internal parts – human diet) we have to measure requirements and supply in terms of kcal of carbohydrates, proteins and fats.

In the same way for energy security:

- When checking the feasibility of energy security against external constraints (context/black-box – "primary energy sources") we have to measure relevant physical quantities in terms of tons of coal, kinetic energy of falling water, cubic meters of natural gas,
- If we want to check the viability in relation to internal constraints (black box/internal parts – "energy carriers") we have to measure relevant quantities in terms of kWh of electricity, MJ of fuels.

In these analysis the required tools and representations are continuously adapting to the task, as the quantification strategy useful to study *feasibility* is not the same as the one to test *viability.* In turn *desirability* demands direct interaction with the social actors carrying legitimate but contrasting normative values (Giampietro, Allen, & Mayumi, 2006).

An additional analytic steps might then combine these findings in a multi-criteria setting (Munda, 2008), or using so called ethical matrices (Mepham, 1996).

QST is our suggested recipe to fight hypocognition in problem structuring (Lakoff, 2004), and contrast this approach to one based on a predefined (set of) model(s), as e.g. it is the case in Economics. Why can't economic truly investigate sustainability?

In our opinion it is very unlikely that by asking the opinion of people about their willingness to pay for ecological services (contingent valuation methods) one can learn about the mechanisms and the processes threatening the surviving of ecological funds. In the same way, the economic viability of a process over a short time scale – e.g. the initial phase of a Ponzi scheme – does not guarantee the long term stability of this solution. Finally, when coming to assessments of desirability of a given policy, the assumption that it is possible to define a common set of preferences for the entire humankind – including youngsters and elderly, women and men, Muslims and Protestants in the same category – is hard to sustain. This implies that whereas economic analysis is very useful for checking the economic viability of economic processes considered at the local scale and to study possible ways to implement policies through incentives, it should be considered inept in the phase of diagnosis of sustainability problems at the large scale.

A desirable feature of QST is indeed that instead of searching for an optimal solution in a given problem space (what economics offers) it strives to enlarge the problem space itself and then map its attribute in terms of feasibility, viability and desirability. This implies that more time is spent on defining the problem and relatively less in populating this with data, indicators and models.

3.2. Quantification in QST

As mentioned above the style of quantification adopted in QST has some characteristic features.

The first is a commitment to a responsible use of quantitative information. For a rich illustration of examples of do's and don'ts in this domain the reader is referred to a recent workshop on the topic organized by the European Commission (EC, 2015, with videos). The list of don'ts include for example refrain from using models that demands as an input data which need to be made up out of thin air, and to refrain from producing digits which do not correspond to the accuracy of the estimate.

The second feature – which is inter alia the one needed to test the salience and relevance of model generated numbers – is to use data and model appraisal strategies developed in the tradition of Post Normal Science (PNS, Funtowicz & Ravetz, 1991, 1992, 1993), which inspires the present special issue on FUTURES. A recent useful review of PNS is in Carrozza, (2014). We mention here NUSAP – a notational system for the management and communication of uncertainty in science for policy (Funtowicz & Ravetz, 1990; van der Sluijs et al., 2005; see also http://www.nusap.net/) and Sensitivity auditing (Saltelli et al., 2013; Saltelli & Funtowicz, 2014), an extension of a purely technical sensitivity analysis to question the assessment generating process in terms of frames, assumptions, interests, agendas, and possible rhetorical uses of the techniques and so on.

Both practices are useful for taming scientific hubris, and for not treating genuine uncertainty or ignorance as if it were amenable to a computable risk, following the key distinction introduced by Knight (1921). In PNS often reference is made to more nuanced taxonomies of uncertainty. One example is from Benessia and de Marchi in this issue. Another example is from Wynne (1992), who distinguishes:

RISK – Know the odds

UNCERTAINTY – Don't know the odds: may know the main parameters. May reduce uncertainty but increase ignorance.

IGNORANCE – Don't know what we don't know. Ignorance increases with increased commitments based on given knowledge.

INDETERMINACY—causal chains or networks open.

For this scholar Science can define a risk, or uncertainties, only by artificially freezing aspects of reality. This process transforms 'evidence' into 'conditional evidence'. Furthermore aspects of a problem might simply be indeterminate. To make an example – e.g. in the field of disposal of nuclear or other hazardous waste material – who can assess the chances that "high quality of maintenance, inspection, operation" will be constantly maintained at different places (wherever) and times (for ever)?

A third key aspect of quantification as advocated in QST is the use of tools from system ecology. For example MuSIASEM or Multi-Scale Integrated Analysis of Societal and Ecosystem Metabolism is a method of accounting based on maintaining coherence across scales and dimensions (e.g. economic, demographic, energetic) of quantitative assessments generated using different metrics (Giampietro & Mayumi, 2000a, 2000b; Giampietro et al. 2014)

3.3. An example

Being QST a new approach we will illustrate its application in a somewhat piecewise fashion, drawing from previous works where some of the ingredients of QST can be identified. We start with the opening the frames, where the best available example is possibly in the field of the persisting controversy surrounding the use of genetically modified organism.

Opposition to GMO food is normally portrayed an unfounded 'scare', and this because GMOs are treated as a nutritional 'risk to health issue' (the given frame). According to upholders of this technology science has given a clear message already by declaring GMO's safe for human consumption. As a result, within this framing of the issue, society or the law should permit its production and consumption.

This framing was found 'irrelevant' in the context of an experimental study where citizens where polled in relation to their GMO concerns (Marris, Wynne, Simmons, & Weldon, 2001). Apparently citizens are largely indifferent or unconcerned about risk posed by GMO to their health. Instead they voiced concern about a rather different set of issues.

- Why do we need GMOs? What are the benefits?

- Who will benefit from their use?
- Who decided that they should be developed and how?
- Why were we not better informed about their use in our food, before their arrival on the market?
- Why are we not given an effective choice about whether or not to buy and consume these products?
- Do regulatory authorities have sufficient powers and resources to effectively counter-balance large companies who wish to develop these products?"

The Marris et al. (2001) study well exemplifies the initial setting of participatory QST approach. With QST this study could be continued, e.g. using MUSIASEM to investigate which real benefit were brought about by GMO.

4. Conclusion

In conclusion we have argued that the evidence based policy paradigm should be revised because in its current use it deploys science to reinforce hypocognition.

We have reviewed existing strand of critique of evidence base policy and taken issue in particular against the use of statistical indicators and mathematical modelling used outside their semantic context as an element of obfuscation and distraction from uncomfortable knowledges. We have suggested an opening of the universe of possible frames which is sensitive to the existence of power relations, of different actors, interests and norms.

We also stressed that shared normative values (frames) are the result of negotiation and shifting of power relations (Lakoff, 2004), and that as a result chosen narratives that resulted useful for in a given historic period to frame existing problems may become useless – when not dangerous or misleading – when the terms of feasibility, viability and desirability have changed.

We suggest that quantitative story telling as described in the present work may represent a less controversial, more socially robust alternative to the present style of quantitative analysis in evidence based policy.

References

Begley, C. G., & Ioannidis, J. P. (2015). Reproducibility in science. Improving the standard for basic and preclinical research. *Circulation Research*, *116*, 116–126. http://dx.doi.org/10.1161/CIRCRESAHA.114.303819.

Begley, C. G., & Lee, M. E. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*, 531–533.

Begley, C. G. (2013). Reproducibility: Six red flags for suspect work. *Nature*, *497*, 433–434.

Benessia, A., Funtowicz, S., Giampietro, M., Guimarães Pereira, A., Ravetz, J., Saltelli, A., et al. (2016). The rightful place of science: Science on the verge. *The consortium for science, policy and outcomes at Arizona State University*.

Boden, R., & Epstein, D. (2006). Managing the research imagination? Globalisation and research in higher education. *Globalisation, Societies and Education*223–236.

Brody, S. (1945). *Bioenergetics and growth; with special reference to the efficiency complex in domestic animals*. Reinhold.

Bush, V. (1945). *Science, the endless frontier*. United States Office of Scientific Research and Development, U.S. Govt. Print Office.

Carrozza, C. (2014). Democratizing expertise and environmental governance: Different approaches to the politics of science and their relevance for policy analysis. *Journal of Environmental Policy & Planning*. http://dx.doi.org/10.1080/1523908x.2014.914894.

Cassidy, J. (2013). *The reinhart and rogoff controversy: A summing up*. The New Yorker. [April 26, 2013 issue].

Collingridge, D., & Reeve, C. (1986). *Science speaks to power: The role of experts in policy making*. London: Frances Pinter.

de Solla Price, D. J. (1963). *Little science big science*. Columbia University Press.

Della Briotta Parolo, P., Kumar Pan, R., Ghosh, R., Huberman, B. A., Kaski, K., & Fortunato, S. (2015). Attention decay in science. *Journal of Informetrics*, *9*, 734–745.

European Commission (2015). *Workshop 'significant digits responsible use of quantitative information'*. https://ec.europa.eu/jrc/en/event/conference/use-quantitative-information.

Fargione, J., Hill, J., Tilman, D., Polasky, S., & Hawthorne, P. (2008). Land clearing and the biofuel carbon debt. *Science*, *319*(5867), 1235–1238.

Funtowicz, S., & Ravetz, J. (1990). *Uncertainty and quality in science for policy*. Dordrecht: Kluwer Academic Publishers.

Funtowicz, S. O., & Ravetz, J. R. (1991). A new scientific methodology for global environmental issues. In R. Costanza (Ed.), *Ecological economics: The science and management of sustainability* (pp. 137–152). New York: Columbia University Press.

Funtowicz, S. O., & Ravetz, J. R. (1992). Three types of risk assessment and the emergence of postnormal science. In S. Krimsky, & D. Golding (Eds.), *Social theories of risk* (pp. 251–273). Westport, CT: Greenwood.

Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures*, *25*(7), 739–755.

Funtowicz, S. O., & Ravetz, J. R. (1994). The worth of a songbird: Ecological economics as a post-normal science. *Ecological Economics*, *10*(3), 197–207.

Funtowicz, S. (2006). What is knowledge assessment? In Â. Guimarães Pereira, S. Guedes Vaz, & S. Tognetti (Eds.), *Interfaces between science and society*, Sheffield: Greenleaf Publishers.

Giampietro, M., & Mayumi, K. (2000a). Multiple-scale integrated assessment of societal metabolism: Introducing the approach. *Population and the Environment*, *22*(2), 109–153.

Giampietro, M., & Mayumi, K. (2000b). Multiple-scale integrated assessments of societal metabolism: Integrating biophysical and economic representations across scales. *Population and the Environment*, *22*(2), 155–210.

Giampietro, M., & Mayumi, K. (2009). *The biofuel delusion: The fallacy behind large-scale agro-biofuel production*. London: Earthscan Research Edition.

Giampietro, M., Allen, T. F. H., & Mayumi, K. (2006). The epistemological predicament associated with purposive quantitative analysis. *Ecological Complexity*, *3*(4), 307–327.

Giampietro, M., Mayumi, K., & Sorman, A. H. (2012). *The metabolic pattern of societies: Where economists fall short*. Routledge.

Giampietro, M., Mayumi, K., & Sorman, A. H. (2013). *Energy analysis for a sustainable future: Multi-scale integrated analysis of societal and ecosystem metabolism*. Abingdon, UK: Routledge.

Giampietro, M., Aspinall, R. J., Ramos-Martin, J., & Bukkens, S. G. F. (Eds.). (2014). *Resource accounting for sustainability assessment: The nexus between energy, food, water and land use*. Routledge. [250 pp].

Grundmann, R. (2009). The role of expertise in governance processes. *Forest Policy and Economics*, *11*, 398–403.

Guimarães Pereira, Â., Funtowicz, S., Eds., (2015), The end of the Cartesian dream, Routledge's series: Explorations in Sustainability and Governance.

Hammersley, M. (2005). Is the evidence-based practice movement doing more good than harm? Reflections on Iain Chalmers' case for research-based policy making and practice. *Evidence and Policy*, *1*(1), 85–100.
Horton, R. (2015). Comment, offline: What is medicine's 5 sigma? *The Lancet*, *385*, 1380.
Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), 696–701.
Ioannidis, J. P. (2014). How to make more published research true. *PLoS Medicine*, *11*(10), e1001747.
Kahan, D. M. (2015). Climate science communication and the measurement problem. *Advances in Political Psychology*, *36*, 1–43.
Knight, F. H. (1921). *Risk, uncertainty, and profit*. New York: Hart, Schaffner & Marx.
Krugman, P. (2000). Thinking about the liquidity trap. *Journal of the Japanese and International Economies*, *14*(4), 221–237.
Krugman, P. (2009). *The return of depression economics*. W. W Norton & Company.
Lakoff, G. (2004). *Don't think of an elephant: Know your values and frame the debate*. Chelsea Green Publishing.
Lakoff, G. (2010). Why it matters how we frame the environment. *Environmental Communication: A Journal of Nature and Culture*, *4*(1), 70–81.
Latour, B. (1991). *Nous n'avons jamais été modernes, Editions La découverte, 1993; We Have Never Been Modern*. Cambridge: Harvard UP.
Leslie, I. (2016). *The sugar conspiracy*. The Guardian. [April 7].
Lyotard, J.-F. (1979). *La Condition postmoderne. Rapport sur le savoir*. Paris: Minuit.
Macilwain, C. (2016). The elephant in the room we can't ignore. *Nature*, *531*, 277.
Marris, C., Wynne, B., Simmons, P., & Weldon, S. (2001). *Final report of the PABE. research project funded by the commission of european communities, contract. number: FAIR CT98-3844 (DG12-SSMI)*. Lancaster: University of Lancaster. [Dec].
Mepham, B. (1996). Ethical analysis of food biotechnologies: An evaluative framework. In B. Mepham (Ed.), *Food ethics* (pp. 101–119). London: Routledge.
Mirowski, P. (1991). *More heat than light. Economics as social physics, physics as nature's economics*. Cambridge: Cambridge University Press.
Mirowski, P. (2011). *Science-mart: Privatizing american science*. Harvard University Press.
Mirowski, P. (2013). *Never let a serious crisis go to waste: How neoliberalism survived the financial meltdown*. Brooklyn: Verso Books.
Munda, G. (2008). *Social multi-criteria evaluation for a sustainable economy*. Berlin: Springer.
Pattee, H. H. (1995). Evolving self-reference: Matter symbols and semantic closure. *Communication and Cognition -Artificial Intelligence*, *12*(1–2), 9–28.
Pearce, W., Wesselink, A., & Colebatch, H. (2014). Evidence and meaning in policy making. *Evidence and Policy*, *10*(2), 161–165.
Phillips, M. (2016). *Science is turning back to the dark ages*. Times. [March 4].
Pierce, A. (2008). *The Queen asks why no one saw the credit crunch coming*. The Telegraph. [5 November].
Pigou, A. C. (Ed.), (1925). *Memorials of alfred marshall*. London: Macmillan.
Porter, T. M. (1995). *Trust in numbers, the pursuit of objectivity in science and public life*. Princeton.
Proctor, R. N., & Schiebinger, L. (Eds.). (2008). *Agnotology, the making and unmaking of ignorance*. Stanford University Press.
Ravetz, J. (1971). *Scientific knowledge and its social problems*. Oxford University Press.
Ravetz, J. R. (1986). Usable knowledge, usable ignorance: Incomplete science with policy implications. In W. C. Clark, & R. Munn (Eds.), *Sustainable development of the biosphere (pp. 415–432)*, New York: IIASA/Cambridge University Press.
Ravetz, J.R. (2015), Descartes and the rediscovery of ignorance, in Guimarães Pereira, and Funtowicz, 2015, p. xv–xviii.
Rayner, S. (2012). Uncomfortable knowledge: The social construction of ignorance in science and environmental policy discourses. *Economy and Society*, *41* (1), 107–125.
Reinert, E. S. (2000). Full circle: Economics from scholasticism through innovation and back into mathematical scholasticism Reflections on a 1769 Price essay:"Why is it that economics so far has gained so few advantages from physics and mathematics?". *Journal of Economic Studies*, *27*(4/5), 364–376.
Rhodium Group (2014). American climate prospectus. Economic risks in the United States. *Prepared as input to the risky business project*. .
Rittel, H., & Webber, M. (1973). *Dilemmas in a general theory of planning*. *Policy sciences*, *Vol. 4*, Amsterdam: Elsevier Scientific Publishing Company, Inc., pp. 155–169. [[Reprinted in N. Cross (ed.), Developments in Design Methodology, J. Wiley & Sons Chichester, 1984, pp. 135–144.].
Saltelli, A., & Funtowicz, S. (2014). When all models are wrong: More stringent quality criteria are needed for models used at the science-policy interface. *Issues in Science and Technology*, *vol. winter*, 79–85.
Saltelli, A., Guimarães Pereira, A., van der Sluijs, J. P., & Funtowicz, S. (2013). What do I make of your Latinorum? Sensitivity auditing of mathematical modelling. *International Journal of Foresight and Innovation Policy*, *9*(2–4), 213–234.
Saltelli, A., Stark, P. B., Becker, W., & Stano, P. (2015). Climate models As economic guides scientific challenge or quixotic quest? *Spring Issue of Science and Technology (IST)*, *XXXI*(3), 79–84.
Saltelli, A., Funtowicz, S., Giampietro, M., Sarewitz, D., Stark, P. B., & van der Sluijs, J. P. (2016). Costing climate is not a science. *Nature*, *532*, 177.
Sarewitz, D. (2000). In R. Frodemen (Ed.), *Science and environmental policy: An excess of objectivity, earth matters: The earth sciences, philosophy, and the claims of community* (pp. 79–98). Prentice Hall.
Schumacher, E. F. (1973). *Small is beautiful economics as if people mattered*. Penguin Perennial.
Searchinger, T., Edwards, R., Mulligan, D., Heimlich, R., & Plevin, R. (2015). Do biofuel policies seek to cut emissions by cutting food? *Science*, *347*(6229), 1420–1422.
Shapin, S., & Schaffer, S. (2011). *Leviathan and the air-pump: Hobbes, Boyle, and the experimental life*. Princeton: Princeton university press.
Siebert, S., Machesky, L. M., & Insall, R. H. (2015). Overflow in science and its implications for trust. *eLife*, *4*, e10825. http://dx.doi.org/10.7554/elife.10825.
Stern, N. (2016). Economics: Current climate models are grossly misleading. *Nature*, *530*, 407–409.
Stiglitz, J. (2011). Rethinking macroeconomics: What failed, and how to repair it. *Journal of the European Economic Association*591–645.
Strassheim, H., & Kettunen, P. (2014). When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms. *Evidence and Policy*, *10*(2), 259–277.
Taleb, N. N. (2012). *Antifragile: Things that gain from disorder*. Random House.
Teicholz, N. (2015). *The big fat surprise: Why butter, meat and cheese belong in a healthy*. Simon & Schuster.
Toulmin, S. (1990). Chicago: University of Chicago Press.
van der Sluijs, J. P., Petersen, A. C., Janssen, P. H. M., Risbey, J. S., & Ravetz, J. R. (2008). Exploring the quality of evidence for complex and contested policy decisions. *Environmental Research Letters*, *3*, 024008. [9 pp.].
van der Sluijs, J., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J., & Risbey, J. (2005). Combining quantitative and qualitative measures of uncertainty in model based environmental assessment: The NUSAP system. *Risk Analysis*, *25*(2), 481–492.
Weinberg, A. M. (1972). Science and trans-science. *Minerva*, *10*, 209–222.
Wilsdon, J. (2014). Evidence-based Union? A new alliance for science advice in Europe. In The Guardian. Available at: http://www.theguardian.com/science/political-science/2014/jun/23/evidence-based-union-a-new-alliance-for-science-advice-in-Europe.
Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*. The University of Chicago Press. [1989 edition].
Wynne, B. (1992). Uncertainty and environmental learning: Reconceiving science and policy in the preventive paradigm. *Global Environmental Change*, *2*(2), 111–127.