Futures 116 (2020) 102509

Contents lists available at ScienceDirect

Futures

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

Essays

Ethics of quantification or quantification of ethics?

Andrea Saltelli

Centre for the Study of the Sciences and the Humanities (SVT), University of Bergen (UIB, Norway), and Open Evidence Research, Universitat Oberta de Catalunya (UOC), Barcelona, Spain

ARTICLE INFO

Keywords:
Ethics of algorithms
Metrics
Statistical modelling
Mathematical modelling
NUSAP
Sensitivity auditing

ABSTRACT

Something can be gained by looking at common ethical features of different instances of quantification. While ethics of algorithms is perceived at present as an urgent issue, similar concerns can easily be associated to the use of metrics, of statistical inference, and of mathematical modelling. By reviewing these common features, strategies of resistance can be imagined to cope with computational dystopias, metrics fixation and numerical abuse.

1. Introduction

Already in the early eighties, sociologist Beck (1992) listed quantitative methods - together with feminism, neo-Marxism and specialization, as relevant anchorage point of the old modernity. Three decades later quantification (metrification, numerification) have greatly expanded in scope, becoming one defining features of the present. Thus, the question of how it is (to be) regulated is relevant and – as discussed here, urgent. Why, though, an overarching ethics of quantification, and not separate ethics for algorithms, for metrics, for statistical or mathematical models? Perhaps, something can be gained by analysing common issues in the various practices of quantification.

One reason is that the explosion of big data analytics and artificial intelligence blurs the difference between data and model: the understanding of the functioning and of the quality of an algorithm (including possible normative biases) becomes impossible without knowledge of the data on which the algorithm has been constructed and calibrated (Brauneis & Goodman, 2018), and when the media discuss about the dangers of Big Data (Redden, 2017), they mostly refer to the use to which these data are put via artificial intelligence systems. The issue of data versus model separability is not limited to algorithms. The existence of model-laden data with data-laden models (Edwards, 1999) – and hence the difficulty to understand one without the other, has already been noted when models and data coevolve in the context e.g. of climate science.

Along the same line, a recent review of five different books on the topic of 'Sociology of Quantification' (Popp Berman & Hirschman, 2018) notes the present blurriness of "quantification", and asks "what qualities are specific to rankings, or indicators, or models, or algorithms?".

A societal reflection has just begun on the ethics of algorithms (Brauneis & Goodman, 2018), where a tension exists between authorities using these tools and the subjects who are on the receiving end. As noted in (Brauneis & Goodman, 2018) authorities keen to deflect critiques may welcome the non-transparency provided by algorithms, but this generates among stakeholders exactly those suspicions and critiques which the authorities wanted to avoid in the first place. Authorities may resort to algorithms to implements rules which - if applied by human subject - would be either legislated or imposed through administrative rulemaking. Instead, big data prediction models can be built and used without policy decisions ever having been the object of a political or a traceable administrative procedure (Brauneis & Goodman, 2018). The range of practices addressed by these algorithmic tools is growing: decisions about hiring and careers; adjudicating the custody of minors; approving or denying credits; choosing which neighbourhoods to patrol; whom to free on parole, and more (O'Neil, 2016).

https://doi.org/10.1016/j.futures.2019.102509
Received 16 October 2019; Received in revised form 24 December 2019; Accepted 30 December 2019
Available online 31 December 2019
0016-3287/

The urgency around an ethics of algorithms owes much to the accelerating pace and impact of the algorithms and artificial intelligence in the new social media (Harari, 2018), with their effects on the democratic process with the so called hacking of elections, (McNamee, 2019), and on the form of the economy associated to the purported emergence of 'surveillance capitalism' (Zuboff, 2019). For others, the same human condition is prey to a new breed of pathologies associated with the new media (Lanier, 2018), while for the jurist Alain Supiot the law is subjugated to the calculation of utilities in a new system of 'governance by numbers' (Supiot, 2007). I argue that while algorithms appear to pose a clear and present danger, other instances of quantification – such as metrics, statistical inference and mathematical modelling, can also have an enduring and damaging influence. The ongoing discussion in statistics, where the same concept of significance is challenged (Amrhein et al., 2019; Gelman, 2019), can be seen together with the so called 'metrics fixation' (Muller, 2018) and with the improper use of mathematical modelling (Pilkey & Pilkey-Jarvis, 2009).

Tolstoy famously noted in his work Ana Karenina that while happy families are all alike, unhappy ones are unhappy in their different ways. Malpractices in the different families of quantification resemble the unhappy families of Tolstoy: an algorithm may embed racial prejudice, a statistical study may be underpowered, a metric may induce goal displacements while a mathematical model may be an instance of mathematical hubris.

Yet, common patterns may exist which justify an overarching ethics of quantification:

- All quantifications are embedded in some scientific discipline, and science is itself a social activity (Ravetz, 1971). The mutated social conditions of the conduct of science, identified as among the root causes of the present scientific irreproducibility crisis (Harris, 2017; Saltelli & Funtowicz, 2017), thus affect all quantifications.
- An example of the above is the Darwinian fitness of bad practices (Smaldino & McElreath, 2016), which tend to reproduce themselves more effectively than the good ones in the present system of incentives (Begley & Ioannidis, 2015). Perverse incentives include the imperative to publish or perish, where production standards are measured by impersonal (sciento-)metrics, and the extreme forms of competition for grant procurements found in the labour market of research (Ruben, 2017). This produces a sort of Gresham's Law of scientific information.
- Along similar lines, science's commoditization (Mirowski, 2011) can influence e.g. cutting statistical corners in medical research (Harris, 2017), while the increasingly mediatization of both science and policy intensifies the use of numbers to score economic advantages or political legitimacy points (Saltelli & Boulanger, 2019).

There is a longstanding relation between societal use of numbers and the existing social-economic order (Shapin & Schaffer, 2011). 'Statistics' was created as an instrument of modern statecraft (Hacking, 1990). Official statistics inevitably reflect the priorities and presuppositions of those whose interests are served by the state (Frankenstein, 1983). Sociology of science points to 'numbers of neoliberalism' (Porter, 2012), whereby a blind faith in numbers underpins the maintenance of the status quo, for example in relation to the existing financial centres of calculation (Ravetz, 2008). For Theodor Porter "funny numbers have given a definite advantage to financial markets", where these numbers are played in the 'theatre of insanity'. Onstage, ostensibly boring numbers are used to shape momentous decision e.g. in relation to financial movements. Offstage the same numbers become the arena for intense struggle and manipulation of the involved parties vying for power and profit. In fact the perils of the financial numbers is as high or possibly higher (Wilmott & Orrell, 2017) than that of the algorithmic ones. I have argued that such elements of insanity emerge frequently in the exercise of quantification, including in statistical and mathematical modelling (Saltelli, 2018a, 2019).

2. What issues for an ethics of quantification?

The case for a new ethics is offered here, in relation to issues scarcely – and never universally, addressed in various domain of quantification.

2.1. The issue of trust

An ethics of quantification is needed because of the symbiotic relationship between quantification and trust (Porter, 1996). A climate of trust favours sensible quantifications, while a climate of regulatory confusion or controversy breeds 'mandated' quantifications, which are less so. Societal trust in numbers plays into the hands of institutions in need of legitimacy, which can gain it by producing a metric and having it accepted by a relevant audience (Porter, 1996).

The controversy surrounding the use of OECD PISA metrics in education (Araujo, Saltelli, & Schnepf, 2017; Meyer & Zahedi, 2014), is a clear illustration of this process: by developing this set of measurements - in a sense a blessing for educationalists and econometricians in need of comparable data across countries - the OECD has gained epistemic authority from national and regional educational authorities. Several scholars have questioned the legitimacy of this process, the metrics' ideological framing, and the distorting effect of goal displacement induced by the PISA tests (Araujo et al., 2017).

2.2. A defence against abuse

What to do when confronted with dubious numbers? Or with practices of metrification or statistical measurement which affect one's own practice or life? Jerry Z. Muller in his Tyranny of Metrics discusses examples of unintended consequences due to the prevalent fixation with the use of metrics (Table 1).

The intended positive function of indicators and benchmarks appears vulnerable to gaming and goal displacement whenever "approval, payment, or some other desired end is made contingent on achieving a quantitative standard" (Porter, 2012). This is known as the Goodhart's (Goodhart, 1981) or Campbell's (Campbell, 1979) law, though an earlier formulation is due to Jerome Ravetz, for whom when the goals of a task are complex, sophisticated, or subtle, then crude systems of measurements can be' gamed' by those persons possessing the skills to execute the tasks properly, who thus manage to achieve their own goals to the detriment of those assigned (Ravetz, 1971; University of California, 2016).

In France a current known as *Stat-Activisme* (Bruno, Didier, & Prévieux, 2014) proposes to 'fight against' as well as 'fight with' numbers, using a variety of strategies against metrification or statistical abuse; these include 'statistical judo' – i.e. gaming the metrics (*a-la-Goodhart*) as an act of self-defence; exposing the hypocrisy or the vacuity of existing metrics, e.g. by denouncing the middle-class bias of an existing consumer price index, and developing a new one in defence of the purchasing power of the worse off; or practicing the classic function of statistics to identify areas of exclusion and neglect. Finally, initiative of self-defence can lead to a totally new system of measurement, as the new barometer for inequality poverty (BIP_{40}), realized and maintained between 2001 and 2007 by the *Réseau d'alerte sur les inégalités*, a French collective of activists in dissent with the numbers of official statistics. With some help from the press (*Le Monde)* the collective managed in 2004 to obtain the attention of the relevant French statistical authorities, to the effect of a greater investment of official statistics itself in the conjoint analysis of both poverty and inequality, realized in partnership with a plurality of social actors. The constructive engagement of the collective with official statistics shows how numbers can be fought with numbers. *Stat-Activisme* takes inspiration from a long tradition of sociology of numbers as discussed in in (Bruno et al., 2014).

2.3. To prevent consequentialism in scientific quantification

Following a purely consequentialist approach to ethics implies that the consequences of one's action are the ultimate basis for a judgment about the rightness of the action. Seen through these lenses a quantification which is particularly effective in communication and mobilizing about a social or environmental threat would be right, in a sense independently from its scientific quality. This could be, for example, the case of the Environmental Footprint. As the time of writing the present article the 2019 edition of the measure – taken up by most media, just informed that in 2019 humans have overshot 100 % of the planet yearly resources by July 29, and that by the end of the year humans activities shall have used 1.75 planets, instead of the one available. The critique that follows might appear to the reader one of those academic disputes which are all the most virulent the less important are the points of contention. In fact, several scholars of different orientation and discipline have noted that while it is evident that man is overexploiting natural resources, the metrification of this into numbers and dates (1.75 planets, July 29) as proposed by the Ecological Footprint does not withstand scrutiny. The details can be read in the exchange of papers written by both proponents and dissenters (Galli et al., 2016; Giampietro & Saltelli, 2014a, 2014b; Goldfinger, Wackernagel, Galli, Lazarus, & Lin, 2014; Blomqvist et al., 2013a, 2013b; Rees & Wackernagel, 2013; Van Den Bergh & Grazi, 2010). Summarizing these critiques – which include the negative judgement of the international commission on the measurement of progress led by Nobel laureates Amartya Sen and Joseph Stiglitz with the French economist Jean-Paul Fitoussi (Stiglitz, Sen, & Fitoussi, 2009) would take much of the space available for the present work. Suffices to say that the 1.75 planets could easily be twenty or two hundred or infinity, if numbers have a meaning, depending on what impacts are measured and how; that the impact of human activity on the planet is too complex to be captured by a single number with three digits accuracy (1.75); and that the concept whereby the earth has a neat yearly budget which humans can use is incredibly optimist against all those impacts which are irreversible – e.g. loss of species, depletion of non-renewable resources, persistent pollutants and so on. The Ecological Footprint is very effective in showing the urgency of action to reduce the pressure on the planet. Why should one be fastidious about the details? A number conveys an irresistible impression of accuracy and allows the setting of target. If humanity could take 5 days away every year from the overshot date (Global Footprint Network, 2019) – as suggested by the proponents, things would go in the right direction. Unfortunately, as agreed by all critics – this measure is particularly ineffective for policy, e.g. in prioritizing what aspects of this pressure to reduce. If there must be an ethics of quantification which is not purely consequentialist, the seduction of the Ecological Footprint should be resisted.

2.4. Who is responsible?

The case of the ecological footprint is not unique. If one were to mention other extremely successful measures of recent times – by impact of society and academia alike, one would likely mention the Shanghai Rankings – known as Academic Ranking of World Universities (ARWU). Neither the ecological footprint nor ARWU were developed by international organizations (such as the OECD, EC, WTO, WEF…), by a statistical office, by a pre-existing disciplinary stronghold or by a think tank. Their fortune owes to the work and the ingenuity of their developers. Though the flaws of ARWU are venial (Paruolo, Saisana, & Saltelli, 2013; Saisana, d'Hombres, & Saltelli, 2011), before the audacious acrobatics of the ecological footprint just discussed, both measures have received pointed criticism from academia and practitioners. Note also that while the ecological footprint scores low on quality, it is not particularly damaging – one would assume, in its societal implications; it reminds humans inhabiting the developed world that they should consume less, and this admonition is couched into an overall non stressing – one might say optimistic, vision of humanity's impact on the planet. To the extent that no actual policy is undertaken to match its scores, - e.g. that no country reduces its food import just to be seen consuming less 'ghost land' (Giampietro & Saltelli, 2014a; Stiglitz et al., 2009), the Ecological Footprint measure makes media satisfied and the publics - in a sense, sedated. The ARWU, while apparently less ambitious in scope, has been rightly noted for its performativity: it forces the university system in the trajectories of a global market for education (Éloire, 2010). Thus, it has been labelled as societally damaging – a least by some observers.

As evident, the media appear to have little interest in technical disquisitions before the effectiveness of an appealing narrative.

Thus, little of the existing criticisms has truly impacted the societal use and take up of these two metrics. This poses the question of who is responsible, when the public intellectuals populating media and academia apparently fail in their role. An interpretation of this conundrum in terms a media system overflowing the specific systems of science, technology and policy is offered in (Saltelli & Boulanger, 2019) using the theoretical implant of Luhmann's social system theory.

2.5. To moderate excesses of optimism about the merits of quantification

The prevailing wisdom is that of a substantial optimism that a quantification is always useful to provide the hard facts on which policies can be adjudicated. In 2018 William Nordhaus, an economist, was one of the winners of the so-called Nobel Prize in Economic Sciences for his work to understand the interactions between society and nature. His "Dynamic Integrated Climate-Economy model" lets the user make different assumptions about environmental policy and see how the consequences of those assumptions may play out over the course of a century. Thus, one might look to assess the impact a carbon-reduction policy will have on a country's economic output from now until 2100 and compare that to a "no policy" scenario or to other policies. An alternative view would be that such quantifications have little relevance, due the large associated uncertainties (Saltelli & d'Hombres, 2010; Saltelli et al., 2015). The prevailing wisdom is likewise illustrated by Cass Sunstein, winner of the Holberg Prize for 2018, also a strong advocate of cost benefit analyses in order to adjudicate policy decisions. When Sunstein said "immersion in the facts often makes value disagreements feel much less relevant" (Matthews, 2018) he was telling us about his preferences rather than asserting a general fact. Following Giandomenico Majone (Majone, 1989), one could inscribe the contribution of Sunstein – including his theory of 'nudging', as an example of Decisionism, a technocratic approach whereby issues can be 'depoliticized' by an economic analysis (Timms, 2019). What is being implied here is not that different instruments of quantification are systematically fallacious, but that excessive faith is being posed in their virtues by the existing system of governance by numbers (Supiot, 2007).

The view that the 'good' can be neatly computed is heir to a long intellectual tradition of quantification, which goes from the *mathématique sociale* of Condorcet (Feldman, 2005) to the utilitarianism of Bentham to the decisionism of post-war bureaucracies (Majone, 1989). A parallel critique has a similarly lengthy tradition. For Langdon Winner – one of the fathers of ecological movement, cost benefit and risk analyses are traps into which environmental activists should not fall (Winner, 1989). For the movement known as post normal science (Funtowicz & Ravetz, 1993) quantifications should be evaluated based on their quality, foremost in relation to their management of uncertainties and their fitness for the intended function (Funtowicz & Ravetz, 1994) – see below the tools NUSAP and sensitivity auditing. The need for vigilance and resistance against malpractices is present in the statistical community (Gigerenzer & Marewski, 2014; Gelman, 2019; Stark & Saltelli, 2018), in PNS (Peters & Besley, 2019), and in the already mentioned current of *Stat-Activisme* (Bruno et al., 2014). In conclusion, the vision defended here clashes with the prevailing wisdom exemplified by Nordhaus and Sunstein, and defined elsewhere (Pereira & Funtowicz, 2015) as a Cartesian dream of prediction and control of man over nature. This minority view benefits nevertheless from the intellectual support of a vast spectrum of scholarship, from ecology to sociology to law, as discussed in the present work.

2.6. For the non-neutrality of techniques

Can it be said that 'The technique is never neutral' (Saltelli, 2018a)? While the statement may appear apodictic, practitioners admit that the choice of the tool can steer the answer in a desired direction.

As noted by Giandomenico Majone (Majone, 1989):

> "In any area of public policy the choice of instruments, far from being a technical exercise that can be safely delegated to the experts, reflects as in a microcosm all the political, moral, and cultural dimensions of policy-making."

Majone (Majone, 1989) also noted that the ostensible distinction between facts and value can be used instrumentally in the policy process, where fact and values cannot always be neatly separated in the making of an argument. A case in point is when the quantification concerns risk. For Ulrich Beck (Beck, 1992).

> Risk determinations are an unrecognized, still underdeveloped symbiosis of the natural and human sciences, of everyday and expert rationality, of interests and facts. They are simultaneously neither simply the one nor only the other.

Beck then goes on to note that in risk determination interests and facts cannot be isolated via specialization, and subjected to standards of rationality. Instead, these determinations require "cooperation across the trenches of disciplines, citizens' groups, factories, administration and politics".

This prescription, identical in many respects to the 'extended peer community' recommended by Funtowicz and Ravetz (Funtowicz & Ravetz, 1993), also includes its own nemesis in that in conflicted issues the cooperation might disintegrate into antagonistic "definitional struggles".

If the example taken from the discussion of risk determinations is taken to show that facts and values cannot always be separated in public policy, this must be true as well for the 'hyper-facts' generated via computer algorithms and mathematical models.

A common case of unethical quantification is when political conflicts and power asymmetries are ignored in favour of reframing the issue as a technical one (Ravetz, 1971). As example, in the field of food ethics (Saltelli & Lo Piano, 2017) better diets and more efficient agricultural systems have been proposed to improve the food scenarios of tomorrow, without mention of the fact that access to food is unequally distributed over the planet between haves and haves-nots, so that the problem is not scarcity, but inequality of provision, and quality of product. As another example, while opposition to genetically modified food and staples has considerably to do with issues of choice, power and political control of an important technology, it is often presented as a unsubstantiated food scare, to be dealt with via better technical communication with the public and by science education (Marris, 2001).

Quantification can be a tool for 'socially constructed ignorance' (Ravetz, 1987), A mathematical model can be effectively used to 'displace' attention away from uncomfortable knowledge. This is achieved when the model – not the issue being modelled – becomes the focus of the attention (Rayner, 2012).

In Economics the non-neutrality of numbers has been the subject of a discussion started by Paul Romer, who coined the term 'Mathiness' to point to the use of mathematics to veil normative stances in growth models (Romer, 2015). For some observers the mathematization of economics is a recurrent problem in the discipline (Reinert, 2000), which is associated by some with a tendency of economists to prefer 'model-land' to the messiness of the real (Smith & A., 2019).

An ethics of quantification needs to be constantly aware of the relationships between knowledge and power (Lyotard, 1979; Ravetz, 1990; Shapin & Schaffer, 2011) – a topic which is too vast to be tackled here, but which is useful as a reminder of the need to consider any quantification in the context of a system of normative frames and power relations, all the more so when quantifications feed into the policy process (Saltelli & Giampietro, 2017; Saltelli, 2018b).

2.7. A need to contextualize quantifications

Any number that does not represent its context and purpose of production runs the risk obfuscating as much as illuminating. In mathematical modelling it is increasingly realized that no validation is possible if the purpose of a model is not specified in advance (Edmonds et al., 2019). Similar recommendations can be found in a recent checklist of good practices in modelling (Jakeman, Letcher, & Norton, 2006). In this respect it is important to note that it is easier to cheat with models - and to get away with it - than it is to cheat with data. The sense of this remark is that while a lot of academic and mediatic attention surrounds statistical malpractices such as P-Hacking (adjusting the data to obtain a usable or publishable inference) and HARKing (adjusting the research hypothesis to the same effect) (Stark & Saltelli, 2018), less attention is given to possible modelling malpractices, such as adjusting the model to the desiderata of the analysts and/or of their clients; see (Saltelli, 2018a, 2019) for a discussion.

A radical version of this precept can be found in (Zyphur & Pierides, 2017) where it is suggested that the quality of a quantitative research must be reframed in terms of "relational validity", understood as the correspondence and alignment of the purpose of the research, in terms of addressing a worldly problem, its orientation, and the way it is performed. If research purposes are defined, then research orientation can be subordinated to them, answering questions such as "for whom is research done?'', "to what end?", "in support of what?" and so forth. This expanded scope prompts thinking broadly in terms of what is to be achieved. Likewise, methods (ways of doing) can thus be aligned with purpose and orientation. The authors note that such an agenda transcends the prevailing disciplinary and positivistic view of facts and reality:

> [...] a theorized singular external world—or, simply, 'reality' in the representation and correspondence narrative—is often understood as being somehow naturally constituted rather than existing as the product of QR [quantified research] practices. In other words, researchers fail to see the rather obvious reality that they coproduce what they propose to merely represent, including populations, variables, statistical parameters, chance or probabilities, and constructs

The authors are clear that such an ambitious agenda is being prescribed for the scope of a single journal (Journal of Business Ethics), and that for such a "vision to be developed a monumental shift in what many QR practitioners care about must occur" (Zyphur & Pierides, 2017).

2.8. To deter quantification hubris

Perhaps nowhere as in the specific case of mathematical modelling one has elements to detect an excessive use of the technique, something which might be called a quantification hubris. I have already mentioned the computation of the cost of climate action/ inaction at the year 2100, which pales when compared to what the authors in (Pilkey & Pilkey-Jarvis, 2009) call "A Million Years of Certainty" relative to a model-based computation of the risk of a nuclear waste disposal one million years into the future.

In environmental modelling a known trade-off is that between modelling bias (or model inaccuracy) and model error, see Fig. 1. This is known as the O'Neill conjecture (see p. 70 in. (Turner & Gardner, 2015)). It implies that neglecting aspects of the phenomenon being modelled may lead to a systematic bias, while being more ambitious in describing the system – making the model more complex, comes at the cost to introducing new parameters, whose estimation entails uncertainties which carry on through the model to increase the error in prediction.

Analogous to this trade-off are known in data science, as under- and over-fitting: if the order of the interpolation polynomial – or the number of nodes in a learning network – is too low, then one may fit poorly the existing (learning) points; if one fits the existing points too well, by increasing e.g. the number of nodes, one will do less well when new points are added. In system analysis the same conundrum is called Zadeh's principle of incompatibility, whereby as complexity increases "precision and significance (or relevance) become almost mutually exclusive characteristics" (Zadeh, 1973). How can the right balance be found? The answer is indeed simple: a proper uncertainty quantification and global sensitivity analysis can map the uncertainty in the model output back to the uncertainty in its inputs, thereby informing on the limits of one's knowledge in relation to the task. It is the impression of this author that most often modellers err on the side of making their model too ambitious, i.e. modelling hubris is more of a problem that model parsimony. In support of this conjecture a classic paper in hydrology notes (Hornberger & Spear, 1981):

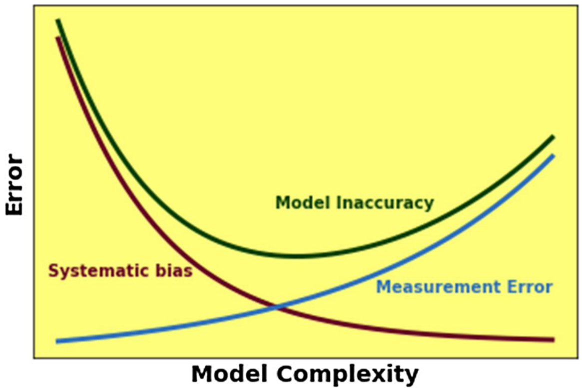

Fig. 1. Model error versus model complexity, adapted from (Turner & Gardner, 2015).

> "[…] most simulation models will be complex, with many parameters, state-variables and non-linear relations. Under the best circumstances, such models have many degrees of freedom and, with judicious fiddling, can be made to produce virtually any desired behaviour, often with both plausible structure and parameter values."

In defence of modellers, the virtues of models simple to the point of unrealism have been extolled by the economist Milton Friedman, who noted that the most significant theories - those most precious for modelling - are based on the most unrealistic of assumptions (Friedman, 1953; Reinert, Endresen, Ianos, & Saltelli, 2016). While this approach has merits, one might ask who decides, though, about what counts as a significant theory. Significant to whom? And for what purpose? Along similar lines economist Joseph Stiglitz argues that the utility of models depends precisely on their acting as 'blinders', which by eliminating from the view the confusion of the real allows economists to investigate in detail the working of the elements left under the microscope. The 'leaving out' is also fraught with normative implications, e.g. when what is left out is a relevant viewpoint on the nature of the problem. These are the same issues just discussed in relation to the work of (Zyphur & Pierides, 2017); see also sensitivity auditing (Saltelli, Guimaraes Pereira, van der Sluijs, & Funtowicz, 2013), discussed below.

2.9. Because of an existing gap

Existing ethical precepts for quantitative research tend to iterate orthodox Mertonian norms of scientific ethos (Merton, 1973), such as universalism, disinterestedness, communalism and organized skepticism, possibly augmented by concerns to protect animal or human subjects of a quantitative study, especially when humans are members of a vulnerable group, see e.g. (The Norwegian Research Ethics Committees (Etikkom) (2015)). Specific norms such as those of official statistics (European Commission & Statistical Office of the European Union, 2017) are absent in other instances of quantification, and this is reflected – in certain fields – in a Feyerabendian 'anything goes', both methodologically and in relation to the consequences of quantification. The case of the OECD-PISA metrics and their unintended consequences have been already mentioned (Araujo et al., 2017; Meyer & Zahedi, 2014). A similar lack of ethical consideration on the consequences of quantitative research in education is lamented for the case of mathematics, in relation to the TIMSS study, and where the situation has led to "Math Wars", especially in the US (Jones, 2000). All cases discussed in the present work, from the Ecological Footprint to the modelling of one million years into the future, bear testimony of the absence of specific norms for responsible quantification.

3. What recipes would be offered by an ethics of quantification?

A quantification hubris should not be replaced with a hubris of recipes. A few general recommendations are offered here.

3.1. A license not-to-quantify

The shortcomings of forced quantification already discussed (Porter, 1996) would suggest that quantification at gunpoint could or should be resisted. In the concluding notes to his Tyranny of Metrics (Muller, 2018) J.Z. Muller notes:

> "Considering all of the above keep in mind at every step that "the best use of metrics may be not to use it at all""

This is not an easy target. In several instances quantifications are compulsory, e.g. grants cannot be secured by a researcher if she does not prove the worth of her work via the existing system of metrics, and nobody in a scientific career – all the more so if young, can realistically imagine to escape the 'metric tide' (Wilsdon, 2016) rolling over scientific production. It is not unusual for a research program to ask the applicants to specify how the impact of a proposal will be measured against which indicators, which evidently condition the nature of the proposal.

3.2. Taming hubris: memento Fig. 1

While in statistics and data analysis methods are available to guard against overfitting, the situation in mathematical modelling is more confused. While a thorough uncertainty quantification and sensitivity analysis would help the analyst to tune the complexity of the model to the quality of the existing data – e.g. to identify the minimum in Fig. 1, these methods are not of general use (Ferretti, Saltelli, & Tarantola, 2016). Often sensitivity analysis is performed ignoring statistical good practices of experimental design and data analysis (Saltelli et al., 2019).

3.3. Make use of the existing disciplinary arrangements

As just mentioned, mathematical modelling could benefit from structure and standards based on statistical principles including a systemic appraisal of model uncertainties and parametric sensitivities. A step in the right direction would be that statistics internalizes these tools into its own syllabi and practices (Saltelli, 2018a, 2019).

3.4. Make quantifications interpretable, conveyable in plain English and context specific; use existing pedigrees

Models – including algorithms, should be made inherently interpretable (Lipton, 2018). This can be achieved via a variety of strategies. Sensitivity analysis could be used to reduce model complexity so that a cut-to-the bone version of the model can be used for an audit or a negotiation. Economists such as Alfred Marshall (Marshall, 1925) and Paul Krugman (Krugman, 2009) recommended using modelling as scaffolding, to be taken out once the edifice of a theory has been built, so that the result can be explained in plain language.

Since all model-knowing is conditional, the conditionality should be made explicit. For key models used in policy, peer review should be enriched to include auditing by an extended peer community involving a plurality of disciplines and interested actors, leading to model pedigrees (Eker, Rovenskaya, Obersteiner, & Langan, 2018; Saltelli, 2018a). As discussed in the first part of this article, any number that does not clarify the context and purpose of its production is incomplete (Edmonds et al., 2019; Jakeman et al., 2006; Zyphur & Pierides, 2017), and this information should be a fundamental ingredient of the quality of a quantification. All these considerations feed into two existing tools for the quality of quantification: NUSAP and sensitivity auditing.

3.5. NUSAP

Addressing virtually all species of quantification, NUSAP (Funtowicz & Ravetz, 1990; van der Sluijs et al., 2005) is a notational system for the management and communication of uncertainty in science for policy, based on five categories for characterizing any quantitative statement. The categories are Numeral, Unit, Spread, Assessment and Pedigree.

- Numeral is the number produced by the quantification, and Unit refers to its units. Spread is an assessment of the error in the value of the Numeral.
- Assessment conveys qualitative judgements – possibly produced by a panel or by an extended peer community - about the quantification. Pedigree, likewise the result of a collective evaluation, looks at the mode of production and of anticipated use of the information.

While the first three are ordinarily met in the quantification practices, assessment and pedigree represent an innovation, whose application presupposes an ‘extended peer community’. An online net (van der Sluijs, 2019) maintains information about applications of the method. NUSAP is suggested in a recent European report by the European Academies’ association of science for policy SAPEA (Science Advice for Policy by European Academies, 2019).

3.6. Sensitivity auditing

Suggested in both the SAPEA report (Science Advice for Policy by European Academies, 2019) and in the European Union guidelines for impact assessment (European Commission, 2009), sensitivity auditing (Saltelli et al., 2013; Saltelli & Funtowicz, 2014) owes much to the perspective and logic of NUSAP, and is tailored to model-based quantification. It proposes seven rules, to be read as a reminder of important aspects of quantification which an analyst might want to investigate. The rules – which can be used both to make sense of an existing quantification or to robustify one against a spectrum of criticisms, are:

- Rule 1: ‘Check against rhetorical use of mathematical modelling’; are results being over-interpreted? Is the model being used ritually or rhetorically?
- Rule 2: ‘Adopt an “assumption hunting” attitude’; this would focus on unearthing possibly implicit assumptions.
- Rule 3: ‘Detect pseudo-science’; this asks whether uncertainty has been downplayed, as discussed above, or amplified, in order to favour an agenda.
- Rule 4: ‘Find sensitive assumptions before these find you’; this is a reminder that before publishing results the analysis of sensitivity should be done and made accessible to researchers.
- Rule 5: ‘Aim for transparency’. This rule echoes present debates on open data, and of the need for a third party to be able to

replicate a given analysis.
- Rule 6: 'Do the right sums'; the analysis should not solve the wrong problem – doing the right sums is more important than doing the sums right. This rule is about asking whether the given quantification is not neglecting important alternatives framings of the issue.
- Rule 7: 'Focus the analysis on the key question answered by the model, exploring holistically the entire space of the assumptions'. An important implication of this rule is that a model cannot be audited for sensitivity once and for all, but needs to be re-audited in the context of each specific application of the model.

Examples of application of sensitivity auditing – of one form or another, are to the OECD-PISA study and the food security case already described in the present work (Araujo et al., 2017; Saltelli & Lo Piano, 2017). Other applications are to the costing of climate change (Saltelli & d'Hombres, 2010), nutrition (Lo Piano & Robinson, 2019), the ecological footprint (Galli et al., 2016; Giampietro & Saltelli, 2014a, 2014b; Goldfinger et al., 2014), and GMO (Saltelli, Giampietro, & Gomiero, 2017).

Used in the context of an ethics of quantification both NUSAP and sensitivity auditing can be used to reveal the normative implications of a quantification. This approach aimed to uncover underlying, unspoken, frames and metaphors has been suggested before in the field of mathematical modelling (Ravetz, 2003) and is applicable to all instances of quantification discussed here.

4. Conclusions

The introduction mentioned that something can be gained by analysing common issues in the various practices of quantification. The analysis has shown common problematic patterns which would lend themselves to the therapies just highlighted.

If the considerations of the present article have some utility, and an ethics of quantification can be usefully conceived, problems would still persist in how its principles could be pursued and disseminated. The idea that numbers as different as

- A consumer price index
- A credit rating
- The price of a bundle of derivatives
- The number of planets needed by humanity in one year
- The gross domestic product to the year 2100
- The numerical risk of an industrial product or practice

need to come under some sort of collective ethical umbrella or even policing may sound as a topic for science fiction. Additionally, not all quantifications involve visible numbers – for example, facial recognition software crunches them just below the surface. It is relevant to note that among the numbers above, the first is regulated in advanced countries by existing codes of conduct (European Commission & Statistical Office of the European Union, 2017). This leads to the question whether the others might be in need of similar arrangements, of the same ambition advocated for quantitative research by (Zyphur & Pierides, 2017). Scholars trained in the use of numbers have the mental habit of questioning their solidity. The same concept is not absent from popular culture. When Homer Simpson says "Oh people can come up with statistics to prove anything, Kent. Forty percent of all people know that" he is simultaneously (ironically) paying tribute both to the existing addiction to numbers and to the underlying suspicion that this amounts to a ritual. The traditional English joke about 'lies, damned lies and statistics' expresses a similar reserve.

One could imagine codes of good conducts for responsible quantification such as those available today for responsible innovation, or for the use of various technologies. While these approaches are not without problems (Strand, 2019), a reflection would be useful (Symposium, Bergen, 2019).

Statistics – of the non-official sort, is itself a discipline, and disposes of the disciplinary fora and leaders to discuss problems such as those addressed here (Saltelli, 2018a). This is less the case for mathematical modelling, metrics, big data and artificial intelligence, though on the latter a considerable ethical work is taking place, including plans of a 'New Deal on Data' (Pentland, 2009). Some thinkers lament (Supiot, 2007) or advocate (Mayer-Schönberger & Ramge, 2018) the advent of an era of government by numbers in order to supplant market as the best adjudicators of everything (Mayer-Schönberger & Ramge, 2018). Whether this would yield dystopia or utopia is the subject of a debate (Bastani, 2019; Morozov, 2019; Mostafa, 2018; Zuboff, 2019). Could or should the committed experts compute society's way to socialism (Bastani, 2019), or save capitalism from its 'surveillance' degeneration (Zuboff, 2019) or strive to control the excesses of financial centres of calculation (Wilmott & Orrell, 2017)? If Alain Supiot is right that the existing governance by number – heir to both Taylorism and Soviet planning, is disrupting all western civilization's familiar legal frameworks - the state, democracy and law itself, and many forms of human solidarity, then we have entered the era of the cybernetic imaginary, which revives the West's age-old dream of grounding social harmony in calculations. Here the goal of governing by just laws is repudiated in favour of the attainment of measurable objectives, where law is subjugated to a computation of utility. This is the 'quantification of ethics' alluded to in the title of the present work.

These are indeed important and urgent questions, beyond the scope of this discussion. The role of an ethics of quantification in this context is all to be discovered and constructed, but to the extent that impending transformations of the existing social orders are under way, such an ethics could play a useful role.

Acknowledgments

Useful suggestions were received by Jerome Ravetz, Bruna De Marchi and by the collective of the Centre for the Study of Sciences and Humanities at the University of Bergen, in particular Gunnar Skirbekk, Thorvald Sirnes, Jeroen van der Sluijs, Marja Sivonen and Scott Bremer.

References

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature, 567*(7748), 305–307. https://doi.org/10.1038/d41586-019-00857-9.

Araujo, L., Saltelli, A., & Schnepf, S. V. (2017). Do PISA data justify PISA-based education policy? *International Journal of Comparative Education and Development, 19*(1), 20–34. https://doi.org/10.1108/IJCED-12-2016-0023.

Bastani, A. (2019). *Fully automated luxury capitalism. A manifesto.* New York: Verso.

Beck, P. U. (1992). *Risk society: Towards a new modernity.* CA): Sage Publications.

Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in science: Improving the standard for basic and preclinical research. *Circulation Research, 116*(1), 116–126. https://doi.org/10.1161/CIRCRESAHA.114.303819.

Blomqvist, L., Brook, B. W., Ellis, E. C., Kareiva, P. M., Nordhaus, T., & Shellenberger, M. (2013a). Does the shoe fit? Real versus imagined ecological footprints. *PLoS Biology, 11*(11), e1001700. https://doi.org/10.1371/journal.pbio.1001700.

Blomqvist, L., Brook, B. W., Ellis, E. C., Kareiva, P. M., Nordhaus, T., & Shellenberger, M. (2013b). The ecological footprint remains a misleading metric of global sustainability. *PLoS Biology, 11*(11), e1001702. https://doi.org/10.1371/journal.pbio.1001702.

Brauneis, R., & Goodman, E. P. (2018). Algorithmic transparency for the Smart City. *Yale Journal of Law & Technology, 20*, 103–176. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3012499.

Bruno, I., Didier, E., & Prévieux, J. (2014). *Stat-activisme. Comment lutter avec des nombres.* Paris: Zones, La Découverte.

Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning, 2*(1), 67–90. https://doi.org/10.1016/0149-7189(79)90048-X.

Edmonds, B., Le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., ... Squazzoni, F. (2019). Different modelling purposes. *Journal of Artificial Societies and Social Simulation, 22*(3), 6. https://doi.org/10.18564/jasss.3993.

Edwards, P. N. (1999). Global climate science, uncertainty and politics: Data-laden models, model-filtered data. *Science As Culture, 8*(4), 437–472.

Eker, S., Rovenskaya, E., Obersteiner, M., & Langan, S. (2018). Practice and perspectives in the validation of resource management models. *Nature Communications, 9*(1), 5359. https://doi.org/10.1038/s41467-018-07811-9.

Éloire, F. (2010). Le classement de Shanghai. Histoire, analyse et critique. *L'Homme et La Société 178*(4), 17. https://doi.org/10.3917/lhs.178.art03.

European Commission, & Statistical Office of the European Union (2017). *European statistics code of practice for the national statistical authorities and Eurostat (EU statistical authority)* (KS-02-18-142). Retrieved fromhttps://ec.europa.eu/eurostat/web/products-catalogues/-/KS-02-18-142.

European Commission (2009). *European commission IMPACT ASSESSMENT GUIDELINES.* Retrieved fromhttp://ec.europa.eu/smart-regulation/impact/commission_guidelines/docs/iag_2009_en.pdf.

Feldman, J. (2005). Condorcet et la mathématique sociale. Enthousiasmes et bémols. *Mathématiques et Sciences Humaines, 172*(4), 7–41.

Ferretti, F., Saltelli, A., & Tarantola, S. (2016). Trends in sensitivity analysis practice in the last decade. *The Science of the Total Environment, 568*, 666–670. https://doi.org/10.1016/j.scitotenv.2016.02.133.

Frankenstein, M. (1983). Critical mathematics education: An application of Paulo Freire's epistemology. *Journal of Education, 165*(4), 315–339. https://doi.org/10.1177/002205748316500403.

Friedman, M. (1953). *Essays in positive economics.* University of Chicago Press.

Funtowicz, S., & Ravetz, J. R. (1990). *Uncertainty and quality in science for policy.* Dordrecht: Kluwerhttps://doi.org/10.1007/978-94-009-0621-1_3.

Funtowicz, S., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures, 25*(7), 739–755. https://doi.org/10.1016/0016-3287(93)90022-L.

Funtowicz, S., & Ravetz, J. R. (1994). The worth of a songbird: Ecological economics as a post-normal science. *Ecological Economics, 10*(3), 197–207. https://doi.org/10.1016/0921-8009(94)90108-2.

Galli, A., Giampietro, M., Goldfinger, S., Lazarus, E., Lin, D., Saltelli, A., ... Müller, F. (2016). Questioning the ecological footprint. *Ecological Indicators, 69*, 224–232. https://doi.org/10.1016/j.ecolind.2016.04.014.

Gelman, A. (2019). *"Retire statistical significance": The discussion.* Retrieved fromhttps://statmodeling.stat.columbia.edu/2019/03/20/retire-statistical-significance-the-discussion/.

Giampietro, M., & Saltelli, A. (2014a). Footprints to nowhere. *Ecological Indicators, 46*, 610–621. https://doi.org/10.1016/j.ecolind.2014.01.030.

Giampietro, M., & Saltelli, A. (2014b). Footworking in circles. *Ecological Indicators, 46*, 260–263. https://doi.org/10.1016/j.ecolind.2014.06.019.

Gigerenzer, G., & Marewski, J. N. (2014). Surrogate science: The idol of a universal method for scientific inference. *Journal of Management*(September), 0149206314547522. https://doi.org/10.1177/0149206314547522.

Global Footprint Network (2019). *Earth overshoot day, 2019.* Retrieved August 13 fromhttps://www.overshootday.org/.

Goldfinger, S., Wackernagel, M., Galli, A., Lazarus, E., & Lin, D. (2014). Footprint facts and fallacies: A response to Giampietro and Saltelli (2014) "Footprints to Nowhere" *Ecological Indicators, 46*, 622–632. https://doi.org/10.1016/J.ECOLIND.2014.04.025.

Goodhart, C. (1981). Problems of monetary management: The U.K. Experience. In A. S. Courakis (Ed.). *Inflation, depression, and economic policy in the West* (pp. 111–146). Rowman & Littlefield.

Hacking, I. (1990). *The taming of chance.* Cambridge University Press, Ed.

Harari, Y. N. (2018). *21 lessons for the 21st century.* Spiegel & Grau.

Harris, R. F. (2017). *Rigor mortis: How sloppy science creates worthless cures, crushes hope, and wastes billions.* Basic Books.

Hornberger, G. M., & Spear, R. C. (1981). An approach to the preliminary analysis of environmental systems. *Journal of Environmental Management, 12*(1).

Jakeman, A. J., Letcher, R. A., & Norton, J. P. (2006). Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software, 21*(5), 602–614.

Jones, K. (2000). A regrettable oversight or a significant omission? Ethical considerations in quantitative research in education. In H. Simons, & R. Usher (Eds.). *Situated ethics inEducational research* (pp. 147–161). London: Routledge.

Krugman, P. R. (2009). *The return of depression economics and the crisis of 2008.* Allen Lane.

Lanier, J. (2018). *Ten arguments for deleting your social media accounts right now.* Henry Holt and Co.

Lipton, Z. C. (2018). The mythos of model interpretability. *ACMQueue, 16*(3), Retrieved from https://queue.acm.org/detail.cfm?id=3241340.

Lo Piano, S., & Robinson, M. (2019). Nutrition and public health economic evaluations under the lenses of post normal science. *Futures, 112*, 102436. https://doi.org/10.1016/J.FUTURES.2019.06.008.

Lyotard, J.-F. (1979). *La condition postmoderne : Rapport sur le savoir.* Éditions Minuit.

Majone, G. (1989). *Evidence, argument, and persuasion in the policy process.* Yale University Press.

Marris, C. (2001). *Final report of the PABE research project funded by the commission of European communities*Contract number: FAIR CT98-3844 (DG 12 - SSMI).

Marshall, A. (1925). In A. C. Pigou (Ed.). *Memorials of Alfred Marshall.* Macmillan & Co..

Matthews, D. (2018). *Can technocracy be saved? An interview with Cass Sunstein. Vox.* (2018, October) Retrieved fromhttps://www.vox.com/future-perfect/2018/10/22/18001014/cass-sunstein-cost-benefit-analysis-technocracy-liberalism.

Mayer-Schönberger, V., & Ramge, T. (2018). *Reinventing capitalism in the age of big data.* Basic Books.
McNamee, R. (2019). *Zucked: Waking up to the Facebook catastrophe.* Penguin Press.
Merton, R. (1973). *The sociology of science: Theoretical and empirical investigations.* University of Chicago Press., Ed.
Meyer, H.-D., & Zahedi, K. (2014). *An open letter: To Andreas Schleicher. May 6*. Retrieved fromThe Guardianhttps://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics.
Mirowski, P. (2011). *Science-mart, privatizing American science.* Harvard University Press.
Morozov, E. (2019). Digital socialism? The calculation debate in the age of big data. *New Left Review,* (116/117), 33–68.
Mostafa, J. (2018). *The revolution will not be automated* (2019, July) Retrieved fromSydney Review of Bookshttps://sydneyreviewofbooks.com/zuboff-bastani/.
Muller, J. Z. (2018). *The tyranny of metrics.* Princeton University Press.
O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy.* Random House Publishing Group.
Paruolo, P., Saisana, M., & Saltelli, A. (2013). Ratings and rankings: voodoo or science? *Journal of the Royal Statistical Society Series A, (Statistics in Society), 176*(3), 609–634. https://doi.org/10.1111/j.1467-985X.2012.01059.x.
Pentland, A. (2009). Reality mining of Mobile communications: Toward a New Deal on data. In S. Dutta, & I. Mia (Eds.). *The global information technology report 2008-2009*. World Economic Forum.
Pereira, A. G., & Funtowicz, S. (2015). *Science, philosophy and sustainability: The end of the Cartesian dream*.
Peters, M. A., & Besley, T. (2019). Citizen science and post-normal science in a post-truth era: Democratising knowledge; socialising responsibility. *Educational Philosophy and Theory,* 1–11. https://doi.org/10.1080/00131857.2019.1577036.
Pilkey, O. H., & Pilkey-Jarvis, L. (2009). *Useless arithmetic: Why environmental scientists can't predict the future.* Columbia University Press.
Popp Berman, E., & Hirschman, D. (2018). The sociology of quantification: Where are we now? *Contemporary Sociology, 47*(3), 257–266.
Porter, T. M. (1996). *Trust in numbers: The pursuit of objectivity in science and public life.* Retrieved fromPrinceton University Presshttps://books.google.es/books?id=oK0QpgVfIN0C.
Porter, T. M. (2012). Funny numbers. *Culture Unbound Journal of Current Cultural Research, 4*, 585–598.
Ravetz, J. R. (1971). *Scientific knowledge and its social problems.* Oxford University Press.
Ravetz, J. R. (1987). Usable knowledge, usable ignorance. *Knowledge, 9*(1), 87–116. https://doi.org/10.1177/107554708700900104.
Ravetz, J. R. (1990). *The merger of knowledge with power : Essays in critical science.* Mansell.
Ravetz, J. R. (2003). Models as metaphors. In W. A. B. Kasemir, J. Jäger, C. Jaeger, T. Gardner Matthew, & C. Clark William (Eds.). *Public participation in sustainability science : A handbook*. Cambridge University Press.
Ravetz, J. R. (2008). *Faith and reason in the mathematics of the credit crunch.* Michaelmas Term 14–16. Retrieved fromThe Oxford Magazine. Eight Weekhttp://www.pantaneto.co.uk/issue35/ravetz.htm.
Rayner, S. (2012). Uncomfortable knowledge: The social construction of ignorance in science and environmental policy discourses. *Economy and Society, 41*(1), 107–125. https://doi.org/10.1080/03085147.2011.637335.
Redden, J. (2017). *Six ways (and counting) that big data systems are harming society. The Conversation.* Retrieved fromhttps://theconversation.com/six-ways-and-counting-that-big-data-systems-are-harming-society-88660.
Rees, W. E., & Wackernagel, M. (2013). The shoe fits, but the footprint is larger than earth. *PLoS Biology, 11*(11), e1001701. https://doi.org/10.1371/journal.pbio.1001701.
Reinert, E. S. (2000). Full circle: Economics from scholasticism through innovation and back into mathematical scholasticism. *Journal of Economic Studies, 27*(4/5), 364–376. https://doi.org/10.1108/01443580010341862.
Reinert, E. S., Endresen, S., Ianos, I., & Saltelli, A. (2016). *Epilogue: The future of economic development between utopias and dystopias. Handbook of alternative theories of economic development*. Edward Elgar Publishing738–786. Retrieved from https://ideas.repec.org/h/elg/eechap/15311_40.html.
Romer, P. (2015). Mathiness in the theory of economic growth. *The American Economic Review, 105*(5), 89–93. https://doi.org/10.1257/aer.p20151066.
Ruben, A. (2017). Another tenure-track scientist bites the dust. *Sciencemag.* https://doi.org/10.1126/science.caredit.a1700056.
Saisana, M., d'Hombres, B., & Saltelli, A. (2011). Rickety numbers: Volatility of university rankings and policy implications. *Research Policy, 40*(1), 165–177. https://doi.org/10.1016/j.respol.2010.09.003.
Saltelli, A. (2018a). *Should statistics rescue mathematical modelling?* ArXiv, arXiv 1712(06457).
Saltelli, A. (2018b). Why science's crisis should not become a political battling ground. *Futures, 104*, 85–90. https://doi.org/10.1016/j.futures.2018.07.006.
Saltelli, A. (2019). Statistical versus mathematical modelling: A short comment. *Nature Communications, 10*, 1–3. https://doi.org/10.1038/s41467-019-11865-8.
Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., ... Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software, 114*, 29–39. https://doi.org/10.1016/J.ENVSOFT.2019.01.012.
Saltelli, A., & Boulanger, P.-M. (2019). Technoscience, policy and the new media. Nexus or vortex? *Futures,* 102491. https://doi.org/10.1016/J.FUTURES.2019.102491.
Saltelli, A., & d'Hombres, B. (2010). Sensitivity analysis didn't help. A practitioner's critique of the Stern review. *Global Environmental Change Part A, 20*(2), 298–302. https://doi.org/10.1016/j.gloenvcha.2009.12.003.
Saltelli, A., & Funtowicz, S. (2014). When all models are wrong. *Issues in Science and Technology, 30*(2), 79–85. Retrieved from http://www.jstor.org/stable/43315849.
Saltelli, A., & Funtowicz, S. (2017). What is science's crisis really about? *Futures, 91*, 5–11.
Saltelli, A., & Giampietro, M. (2017). What is wrong with evidence based policy, and how can it be improved? *Futures, 91*, 62–71. https://doi.org/10.1016/j.futures.2016.11.012.
Saltelli, A., Giampietro, M., & Gomiero, T. (2017). Forcing consensus is bad for science and society. *The Conversation*(May 12).
Saltelli, A., Guimaraes Pereira, Â., van der Sluijs, J. P., & Funtowicz, S. (2013). What do I make of your latinorumc Sensitivity auditing of mathematical modelling. *International Journal of Foresight and Innovation Policy, 9*(2/3/4), 213–234. https://doi.org/10.1504/IJFIP.2013.058610.
Saltelli, A., & Lo Piano, S. (2017). Problematic quantifications: A critical appraisal of scenario making for a global "sustainable" food production. *Food Ethics, 1*(2), 173–179.
Saltelli, A., Stark, P. B., Becker, W., & Stano, P. (2015). Climate Models as Economic Guides - Scientific Challenge of Quixotic Quest? *Issues in Science and Technology, 31*(3), 1–8.
Science Advice for Policy by European Academies (2019). *Making sense of science for policy under conditions of complexity and uncertainty*. Berlin. Retrieved fromhttps://www.sapea.info/topics/making-sense-of-science/.
Shapin, S., & Schaffer, S. (2011). *Leviathan and the air-pump: Hobbes, Boyle, and the experimental life: With a new introduction by the authors.* Princeton University Press.
Smaldino, P. E., & McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science, 3*(160384).
Smith, E. L., & A, T. L. (2019). Escape from model-land. *Economics, the Open Access, Open Assessment Journal, 23*(March 08).
Stark, P. B., & Saltelli, A. (2018). Cargo-cult statistics and scientific crisis. *Significance, 15*(4), 40–43. Retrieved from https://www.significancemagazine.com/2-uncategorised/593-cargo-cult-statistics-and-scientific-crisis.
Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2009). Report by the commission on the measurement of economic performance and social progress. *Sustainable Development, 12*, 292. https://doi.org/10.2139/ssrn.1714428.
Strand, R. (2019). Striving for reflexive science. *Fteval Journal for Research and Technology Policy Evaluation, 48*, 58–63.
Supiot, A. (2007). *Governance by numbers: The making of a legal model of allegiance.* Oxford University Press.
Symposium: Ethics of quantification, Centre for the Study of the Sciences and the Humanities, University of Bergen, December 2019. (2019). Retrieved July 3, 2019, from https://www.uib.no/en/svt/127044/ethics-quantification.
The Norwegian Research Ethics Committees (Etikkom) (2015). *Quantitative methods.* Retrieved September 14, 2019, fromhttps://www.etikkom.no/en/library/introduction/methods-and-approaches/quantitative-methods/.
Timms, A. (2019). *The sameness of Cass Sunstein | the New Republic* (2019, June) Retrieved fromNew Republichttps://newrepublic.com/article/154236/sameness-cass-

sunstein.
Turner, M. G., & Gardner, R. H. (2015). *Introduction to models. Landscape ecology in theory and practice*. New York, NY: Springer New York63–95. https://doi.org/10.1007/978-1-4939-2794-4_3.

University of California, D (2016). *Gaming metrics: Innovation & surveillance in academic misconduct.* Retrieved August 22, 2019, fromhttps://www.library.ucdavis.edu/news/gaming-metrics-innovation-surveillance-in-academic-misconduct/.

Van Den Bergh, J., & Grazi, F. (2010). On the policy relevance of ecological footprints. *Environmental Science & Technology, 44*(13), 4843–4844. https://doi.org/10.1021/es1003582.

van der Sluijs, J. P. (2019). *NUSAP.* Retrieved fromhttp://www.nusap.net/.

van der Sluijs, J. P., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J. R., & Risbey, J. (2005). Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: The NUSAP system. *Risk Analysis, 25*(2), 481–492. https://doi.org/10.1111/j.1539-6924.2005.00604.x.

Wilmott, P., & Orrell, D. (2017). *The money formula.* Wiley & Sons.

Wilsdon, J. (2016). *The metric tide: The independent review of the role of metrics in research assessment and management.* Retrieved fromSage Publications, Ltd.https://uk.sagepub.com/en-gb/eur/the-metric-tide/book251812.

Winner, L. (1989). *The whale and the reactor: A search for limits in an age of high technology.* University of Chicago Press.

Zadeh, L. (1973). Outline of a New Approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics, 3*(1), 28–44.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power.* PublicAffairs.

Zyphur, M. J., & Pierides, D. C. (2017). Is quantitative research ethical? Tools for ethically practicing, evaluating, and using quantitative research. *Journal of Business Ethics, 143*(1), 1–16. https://doi.org/10.1007/s10551-017-3549-8.