What's Observed in a Rating?
Rankings as Orientation in the Face of Uncertainty

Elena Esposito
University of Bielefeld
and University of Bologna

and

David Stark
Columbia University
and University of Warwick

Keywords: ratings, rankings, valuation, second-order observation

What's Observed in a Rating?
Rankings as Orientation in the Face of Uncertainty

Elena Esposito and David Stark

Abstract: Ratings and rankings are criticized for being simplistic, obscurantist, inaccurate, and subjective, yet they are becoming an increasingly influential social form. We elaborate the criticisms of ratings and rankings in various fields but go on to argue that analysis should shift its target. The problem that ratings deal with is not observation of an independent world. Instead, the challenge they face is the circularity of second-order observation in which observations must take into account the observations of others. To this purpose they function well enough not because they mirror how things are but because they offer a highly visible reference point to which others are attentive and thereby provide an orientation to navigate uncertainty. The concluding section places the problem of ratings and rankings in a broader historical perspective contrasting the ranked society to the society of rankings. Responding to uncertainty, ratings and rankings perpetuate rather than eliminate anxiety.

INTRODUCTION

They are ubiquitous:

The 100 Best Colleges in the US

The Top 100 Hospitals

The Top 100 Restaurants in Edmonton

Top 100 Songs of the 1960s

Top 100 Movies of all Time

Top 100 EBooks Yesterday

The World's Top 10 Strongest Democracies

The Top 50 Universities under 50 years old

The Top 100 Pizza Toppings.

We are surrounded by ratings and rankings. Banks are rated for financial strength, nations are rated for their level of corruption (Cooley and Snyder 2015), celebrities are ranked by their bankability, and prisons are ranked according to performance standards (Mennicken 2016). Professors are rated by their students, their articles hierarchized by

citation counts, their journals ranked by impact factors, and their disciplines scored by their relative sexiness (Felton et al 2008). Wines can be ranked by their Parker score (Chauvin 2016), your credit rating expressed by your FICO score (Poon 2007), and the impact of your online presence in the social media world measured by your Klout score (Gerlitz and Lury 2014).

Some ratings and rankings have been around for awhile. The controversial financial ratings agencies, for example, have been in operation for about a century (Langohr and Langohr 2008, pp. 377 ff.). The first blue book Michelin guide was issued in August 1900. Edward O'Brien began editing his annual compilation of *The Best American Short Stories* in 1915. The *New York Daily News* first began grading movies on a scale of zero to three stars in July 1928. But if the ranked list has a long pedigree (going back at least to Moses's stone tablet list of the Top Ten Prohibitions), it is also the case that the rating frenzy is a relatively recent phenomenon. The National Academy of Sciences rankings of graduate programs in engineering and natural sciences began in 1925, but was primarily intended for academics (Stuart 1995). In fact, the craze over college and university rankings really began only after *US News and World Report* launched its rankings in 1983 (Espeland and Sauder 2007; Musselin 2010). The spread of ratings and rankings in many, very different fields had a sharp acceleration only in the most recent decades (Kornberger and Carter 2010: 33; Grewal at al. 2008). For example, it is only since the 1990s that the financial ratings agencies have had a recognized and explicit role in the regulation of the banking system. At the international level, 53 of the prior-mentioned 82 public ratings of nation states identified by Cooley and Snyder (2015) appeared since 2001.

While they are becoming more and more pervasive, ratings and rankings are hugely controversial and very much criticized in all areas. Rating Agencies have been charged with a good part of the responsibility for the recent financial crisis: they evaluated bad, late, and evaluated the wrong things. University rankings still find great hostility among academics, who consider them stupid, superficial, self-interested, harmful, and unable to grasp the aspects really significant for research and teaching (Espeland and Sauder 2007). The credibility of restaurants guides is doubted both by chefs and by customers (Ayeh et al 2013) and their effects are ambiguous: the reference

to reviews and rankings risks becoming a substitute for live interaction and thereby killing adventure in traveling (Kugel 2018).

People know it and complain. However, these tools continue to proliferate and become increasingly influential as the web – which seems to "think" in the form of lists (Poole 2013, Eco 2009) – further contributes to the explosive growth of ratings and rankings as well as denunciations of them as misguided. Rating Agencies still find buyers for their reviews, which continue to be discussed in the press and in the mass media in general – even when their indications are not followed by traders (for example the downgrading of France in November 2011, which was basically ignored by markets). College rankings, criticized and often openly opposed by academics, have spread so much that it is very difficult to ignore them – even at the expense of the quality of teaching and research. Parker's rating of wines serves as a general reference for everyone, also for those who oppose the trend towards increasing "parkerized" wines (Chauvin 2014: 90). As Cooley and Snyder (2015) ask provocatively, "If ratings are so fraught, why are they so popular?"

This paper addresses that question, examining the curious situation around the explosion of ratings and rankings from a distinctive perspective: We argue that the features for which ratings and rankings are rightly criticized are the same mechanisms that underlie their effectiveness. There are, of course, wide differences between disparate phenomena like lists on the web, wine ratings, or university rankings.[1] It is not by chance, however, that they are spreading at the same time and share many formal features.[2] To understand these phenomena and their presuppositions beyond the immediate differences we need a broader perspective that considers them jointly and refers to the underlying question they all address: social complexity and efforts to cope with it.[3] This requires abandoning the idea of reality mirroring. As we shall see, the

---

[1] Not all lists are rankings; many are not even ratings. And not all ratings are rankings. Ratings evaluate but do not necessarily compare their objects. Rankings compare the items and establish a hierarchy, while pure lists do not evaluate and do not compare (Esposito 2017).

[2] As Espeland and Sauder (2016: 5) observe, the patterns of evaluation and response created by university rankings are very similar in different contexts subject to accountability metrics.

[3] Note on terminology: To avoid unnecessary repetition of "ratings and rankings," except for passages when we are explicitly making distinctions between ratings and rankings, we use the

problem that ratings deal with is not observation of an independent world (for which they are inevitably flawed and inadequate) but the circularity of second-order observation in which observations must take into account the observations of others. To this purpose they function well enough, *not because they inform us about how things are but because they provide an orientation about what others observe*. It is in regard to problems of second-order observation that ratings and rankings are used, offering a reference in a contingent world with its horizon of uncertainty.

In the section that follows, we summarize the criticisms of ratings and rankings, grouping these along four major lines: 1) Ratings simplify, 2) they are intransparent, 3) they fail as forecasts, and 4) they are not objective. In the subsequent section, we introduce an alternative perspective from which to analyze ratings and rankings, elaborating that framework in the next section by revisiting these four features. In the concluding section we place the problem of ratings and rankings in a broader historical perspective contrasting the ranked society to the society of rankings.

<div align="center">BERATING THE RANKINGS</div>

Ratings and rankings are criticized in every field, from universities to hospitals, from wines to finance. The complaints are remarkably similar across fields and can be grouped along four basic contours.[4]

1. Ratings and rankings simplify.

Ratings offer a one-dimensional quantitative measure (Karpik 2007: 331) that brutally reduces the qualitative complexity of phenomena and disregards their local, cultural, and social contexts (Cooley and Snyder 2015; Jeacle and Carter 2011: 297;

---

terms interchangeably in this paper. Our object of study is the general social form of the evaluative list.

[4] Espeland and Sauder (2016: ch.2) also identify four mechanisms at the base of the power of rankings: commensuration, self-fulfilling prophecies, narrative, and reverse engineering. While their analysis partly coincides with ours in underlining the performative effects of evaluation (discussed below), Espeland and Sauder refer their four mechanisms to the "reactivity" activated by rankings. For us instead, as we will argue, reactivity itself is a consequence of second-order observation. What interests us here is to highlight the characteristics of rankings that make such observation possible and to identify the first-order perspective from which criticisms are formulated.

Kornberger and Carter 2010: 328). In the financial field they fail to take into account the many dimensions of risk (Levich et al. 2002: 13); in the oenological field they reduce taste to a variety of sensory measures; and in the academic field they flatten the different approaches of colleges onto a single plane, erasing distinctive options (for example, moral or religious choices) and additional offerings (tutoring, counseling, basic research) that do not fit the ratings' evaluative metrics. The ranking of prisons in the UK aggregates in a single number very disparate dimensions such as number of escapes, educational effectiveness, resource management, and telephone answering time (Mennicken 2016), even though some of them can be at odds with each other – an effective educational program for inmates, for example, might not be compatible with a reduction in the number of escapes.

Quantification integrates information but, in so doing, reduces, simplifies, and decontextualizes knowledge, thereby rendering many components irrelevant while imposing a singular form to all the rest (Espeland and Sauder 2007: 17). The problem with a rating begins from the moment it is expressed in a number, one that is larger or smaller than the one attributed to the other evaluated objects. It is this number form that facilitates comparison. Or, stated more forcely, as some critics suggest (Espeland and Stevens 2008), it is this number form that *forces* comparison.[5]

The simplicity of quantification masks the fact that ratings are built around disparate and incommensurable principles of evaluation. It must be such because every scientific article, every organization, every wine, strictly speaking every object, is unique and different from every other one.[6] In many cases, ratings are created without the intention of comparison. In such cases, ratings begin with scores. In classic guides like Michelin, originally the various features of restaurants were dealt with separately:

---

[5] See Stark (2011) on the tendency for ratings to become rankings. There are exceptions, as in the case of the rating of prisons in England (Mennicken 2016), but they refer to ratings used for administrative purposes rather then for public orientation.

[6] In a fascinating study, Kreiner (2010) examines these problems in the case of architectural competitions. The conceptual difficulties spring from the fact that different projects cannot be compared because they do not offer answers to the same question. The various teams of designers in the competition each interpret the task in their own way. Indeed, they must do so because this is part of the creativity of their performance. The results are then different answers to different questions, and these are what rankings (as the result of the competition) should order.

quality of raw materials and preparation, for example, but also originality, atmosphere, view, and many other factors that cannot be sensibly aggregated into a single measure (Karpik 2007: 113ff.). What these raw scores offer is a multiplicity of singular judgments. But, once these are available, they tend, almost inevitably, to be put into a comparative mode that yields a rank order (Guyer 2010: 124). From the 1930s, the Michelin guide moved to the aggregate classification in one, two, or three stars.

Thus, comparison emerges even if it was not the intent of those who produced the numbers. From multifaceted judgments along many dimensions, the simplified results are the number of stars. From the multiplicity of scorings assessing various aspects of the object, the institution, or the performance, there is now a single score which can easily translate into a position within a ranking. In the current debate, it is this transformation into a ranking that is the principle villain.[7]

The central problem, therefore, is that *singularities are reduced to comparabilities*. A ranking displaces complexity, yielding a hierarchy in which each element has a lower place than the previous one and a higher place than the next one. The hierarchy monopolizes attention. The users of rankings look at who's up and who's down, not at what it is. The ranking describes the mutual relations between a number of entities, and not the performance of each of them (Stuart 1995) - much to the frustration of the entities involved, who can happen to make significant individual progress without changing at all the position in the rankings (because in the meantime the others also moved) (Cooley and Snyder 2015; Schulz et al. 2001: 37).

Several problems derive from this reduction of singularities to comparabilities. Among these, critics bemoan that comparability leads to competition. What was supposed to be a descriptive measure to provide information to the user about the object (e.g., the quality of a university or a restaurant) tends to become an assessment

---

[7] Let's take for example the recent alarm on university rankings, which even attracted the attention of the White House (Shear 2014). In the face of the many dysfunctional consequences of the rankings of colleges, it has been proposed to replace the (uncontrollable, suspect, generic) rankings with a system of ratings, which offer a series of evaluations on various aspects of each institution, without aggregating them in an overall order. Ratings should be more stable and controllable, "driven by a commitment to quality, not by an obsession with a numerical standard" (Policano 2005: 32).

that puts it in competition with other items. Competition is not a "natural fact" that ratings merely reflect, but rather a consequence of the assessment. Entities such as cities, for example, before the production of league tables that assess and implicitly compare them, were not "better" or "worse" than the others, but were just as they were – with their histories, their specificities, and their incomparable features (Kornberger and Carter 2010). The production of league tables of cities, starting from the 1980s, generated debate and competition among them. This socially constructed competition changed how cities were observed (both internally and externally) and produced a series of consequences in policy and resource management.

2. Ratings and rankings are intransparent.

Ratings are black boxes. Who relies on ratings, as in general on the various forms of audit, relies on an opinion (Power 1997: 175; Langohr and Langohr 2008), without really knowing upon what it depends because the system of ratings – ranging from the structure of evaluating entities, to their evaluation criteria, to their metrics, procedures, and algorithms – is substantially opaque. The ownership and internal organizational structure of the evaluating entities, critics claim, are obscure. In some cases (for example, in financial firms but also for universities or restaurants), the metrics, procedures, and algorithms used in the evaluation are officially and explicitly kept confidential, allegedly to avoid manipulation by the rated entities. Where ratings rely on experts, criteria are not disclosed lest they lose their privileged status (Blank 2007: 8).[8] Moreover, in many cases the work of evaluators requires the availability of confidential information provided by companies, such as budgets, forecasts, contingent risks analyses, and information relating to new financings, acquisitions, dispositions and restructurings – and this confidentiality is respected.

The lack of transparency in the structures and processes of the rating system is matched by a similar lack of transparency in its legitimation. The rating of educational institutions, for example, is clearly a matter of public concern yet it is predominantly

---

[8] But the counter-trend, as in TripAdvisor, towards user-based reviews, where giving up control of the sources is the source of credibility (Jeacle and Carter 2011), keeps evaluation criteria just as obscure.

conducted by for-profit private rating organizations based in the media world with dubious academic legitimacy. This situation is especially troubling in fields where rating and ranking organizations have emerged in recent decades as new regulatory actors with considerable power (Cooley and Snyder 2015). In the financial sector, for example, since the Basel 2 Agreement the assessment of credit risk used to calculate minimum capital requirements for banks is to be carried out by acknowledged financial ratings agencies (Partnoy 2002; White 2002). These agencies produce a public good, yet their operation as private entities is largely protected from the gaze of the public. They have not been elected, were not chosen, were not even nominated, and their legitimacy remains intransparent, outside the codified procedures of public debate.

This lack of transparency is even more suspect because of the peculiar dependencies linking rating organizations to the rated entities. In many cases the evaluators are tied to the evaluated. These links are sometime direct as, for example, with the financial ratings agencies whose work is paid and judgments are purchased by the rated entities. Often, these links are indirect, as when evaluators and evaluated come from the same milieu (restaurant reviewers know the chefs, wine raters have long-standing ties to vintners, and so on). Moreover in complex situations, such as the current financial arena, raters inevitably act as consultants and affect firms' decisions, indicating for example how they would rate them if re-organized in particular ways (Sinclair 2010: 8). How can we trust the neutrality of their judgment?

3. Ratings and rankings fail as forecast

Ratings are frequently consulted in hopes of obtaining a description of the world that lies ahead. But, as Yogi Berra noted, "It's tough to make predictions, especially about the future." The future predicted by the rating can be observed only ex-post, and in social life the problem becomes even more complicated because even in retrospect one cannot know whether success or failure really indicate that the assessment was correct or false. Was the rating confirmed because it was an accurate judgment or because the object adjusted to the evaluation (e.g., markets modified as a result of the information

provided by ratings)?[9] If the prediction failed, was it because it was wrong or because, being correct and credible, it was followed by everyone, thereby changing the conditions of the world (the excellent family restaurant became mediocre and standardized as a consequence of the recommendation in the guide)?

For these and other reasons some critics suggest that in social life ratings have little informational value about the future and, in fact, that it is difficult to figure out if and how they have any informational value at all. Rating assessments notoriously lag behind. Ratings often retrospectively model what has already happened (Cooley and Snyder 2015) and construct an image of the future that is merely a forward projection of data from the past. Financial ratings run after the market, which anticipates the majority of changes in ratings (Partnoy 2002). Research shows that sovereign credit ratings systematically fail to anticipate crises in emerging countries (Reinhart 2002). The measures ratify the changes already underway and produce pro-cyclical trends.

The future is made of surprises, but the data available in the present allow at best to protect against expected surprises. Yet in social situations it is the unexpected surprises (the low frequency high impact events) – which cannot be derived from current trends – that are the really interesting ones. A financial analyst would like to know which unpredicted event can endanger the reserve capital of the bank; a future student would like to know which skills will boom on the future labor market, i.e., not the ones already emerging today. But this is precisely the information that ratings are not able to provide.

4. Ratings and rankings are not objective

Ratings, goes the prevalent view, should provide objective information, available for everyone regardless of the point of view and the interests of the observer. The more objective, the better they can fulfill their informational function. Their function is to overcome the obstacles standing between the user and the necessary information in the proper form useful to make a decision. Such obstacles can be information asymmetries that block the user (as outsider) from essential insider knowledge or lack of time and competence to adequately evaluate available information.

---

[9] Ratings tend to activate a process of *reactivity* whereby the evaluated institution (law schools, for example) adjust their practices to conform to the criteria of the rating agencies (Espeland and Sauder 2007: 6ff.; 2016, ch.2).

Critics claim, however, that when ratings are subjected to scrutiny, one quickly discovers that their judgments lack objectivity. As evidence for the lack of objectivity, critics point to the fact that movements in placements on the lists of rankings inexplicably change too quickly. As such a case, given the stickiness and complexity of metropolitan transformations, how can a city like Los Angeles move in a single year from 10th to 17th place on a ranking of cities (Kornberger and Carter 2010: 338)? It is implausible that the quality of life in Los Angeles could be altered so much from one year to the next. But, for other critics, it is instead the lack of change that indicates the presence of bias. Stickiness in this case affects the placement in the rankings, with the entities at the top keeping their positions over time, even if the world, the criteria and the evaluation methods change (Schulz et al. 2001). In the 1999-2006 list of *US News and World Report* top 50 universities 47 institutions were always present, while Harvard, Princeton, Stanford and Yale were in the top five throughout (Grewal et al. 2008). Can such a rigid classification be trusted?

When prompted to defend their objectivity, evaluators attempt to dispel any idea that their work is judgmental or arbitrary; and they are quick to claim that they follow standardized and tightly regulated processes (Pollock and D'Adderio 2012: 575).[10] But in the world of ratings, as in any domain of social life, reliability (in the sense of commitment to procedures and compliance with rules) does not imply general validity (Cooley and Snyder 2015). Take, for example, the status of "friendly" or "affordable" in the evaluation of cities, or "courtesy" and "elegance" in restaurants. These and similar notions inevitably depend on the perspective of the group of persons the rating intends to address. And even when they intend to measure objective characteristics, such as the "locational advantages" offered by large cities for those in business (e.g., the number of MBA programs, Google hits, patent applications, daily newspapers, and the like), the indicators depend on a choice of the evaluator, and do not necessarily actually measure the entity you want to know. The indices, on the whole, rather than on the world they should measure inform on those who measure and on those who are measured – on how

---

[10] Rating agencies themselves emphatically repeat that they offer are no statements of fact (Langohr and Langohr 2008: 17, 474) – but also claim rather contradictorily that they offer opinions that "objectively measure the credit risk of the issuers business and its debt financing and resolve information asymmetry about it" (89).

they operate and how they react to the measurement (Kornberger and Carter 2010: 335ff.).

RANKIIGS AS REFERENCE POINTS IN A WORLD OF UNCERTAINTY

Ratings and rankings, critics contend and we agree, are simplistic, obscurantist, inaccurate, and subjective. Unfaithful to their promise as neutral sources of valuable information, they nonetheless circulate in widening spheres and proliferate in increasing numbers. Untrusted, they are nevertheless consulted. Whether we pay for them with our money or, when freely available, pay for them with our allocated time, we pay attention.

To understand why we pay attention – that is, to grasp the role and functioning of ratings and rankings – we will need to change perspective, starting from assumptions about the challenges facing the consumer or decision maker who consults a ranking. By one way of thinking, the user seeks information about the qualities of the products, services, or institutions being considered. In a world in which more and more options are competing along more and more dimensions, how to decide which is the more valuable? The ranking, so it would seem, offers such information in a straightforward format. And the better it represents the underlying reality and the better it demarcates the more valuable from the less valuable, the more valuable it is to the decision-making user.

The problem with this view it that it vastly underestimates the complexity of the observational character of the problem of "what's valuable?" in our data-rich society. Complexity means that there are more possibilities than can be taken into account, and that you know it (Morin 1985; Luhmann 1984: 45 ff.). Any data could be connected with other data, and could thereby look different – whether by someone else somewhere else who is dealing with it in another way or by us ourselves in the future. Complexity dramatically increases contingency (the possibility to be otherwise). It amplifies uncertainty and compounds the difficulty of making decisions. How do I decide that the available information is reliable if the more I collect information the more this inadequacy will be obvious to me? And how do I decide, knowing that important information will be produced as a *result* of my decision and therefore cannot be known before I act? Nevertheless I must decide, because not deciding also has consequences.

These are the typical dilemmas of our risk society (Beck 1986; Luhmann 1991), which do not only affect big public choices (nuclear power, genetically modified organisms, global pollution) but also small and middle range daily decisions: to which school should I enroll my children, how do I organize my pension plan, which wine should I buy for a dinner with my colleagues? In all these cases we do not have all the information necessary to decide with certainty, first because this information depends on what others think and do: the right school is the one whose name will be best evaluated by employers, the right wine is the one that will make a good impression on my guests.

This is particularly elusive because the others also observe, and what they observe depends in turn on what I and the other observers do. This inevitable circularity makes the issue unstable and difficult to manage. This circularity creeps in even in the most personal aspects as taste, for example in the by now extremely complex selection of quality wines. In front of the multiplicity of skills that one should have, often the wine that we know to be considered the best is the one we like best: taste is the consequence, not the premise of the evaluation (Hennion 2015; Karpik 2007: 302). Of course we can always disagree, but in order to do it we must have an orientation – we must disagree with something (Chauvin 2014).

How do ratings and rankings facilitate action amidst such circuits of uncertainty? We claim that they are effective under conditions of high complexity because they shift the reference from the world to observers – i.e., from first-order observation to second-order observation. The distinction of first-order and second-order observation has been introduced by Heinz von Foerster (1981) to distinguish the condition in which observers focus on objective data from the one in which they turn to the perspective of other observers. As von Foerster argues, in the shift to second-order observation reality does not disappear, but does not coincide with objectivity any more. Reality is not the starting point, it is the result of observation, produced by the reciprocal reference of observers to the perspective of others. This is the most reliable reference in a world that has become too complex for univocal determinations. Even if not objective, this multiple reality is by no means arbitrary. What observers observe is contingent (in the sense that it would be different from another perspective) but cannot be changed at will. Once a reference has

been chosen and shared, the perspective is binding, effectively excluding any arbitrariness.

Second-order observation does not imply knowing what relevant others are thinking. This is an impossible fantasy, only increasing uncertainty. Bypassing this uncertainty, ratings and rankings offer a reference about which one can be relatively certain. I cannot know the thoughts of relevant others, and I cannot know all of the data and information to which they are exposed, but I can know that there is strong likelihood that others observe the ranking. I need not trust in the reliability of the ranking to have reasonable confidence that others have consulted it. The ranking is an opinion; but not just any old opinion. It is one that I pay attention to because others pay attention to it. Rankings are important because people consider them important, and everyone knows it.

The "opinion" of the ranking is an opinion that everyone observes and can observe others observing. The visible public character of the ranking makes it a common reference point (Langohr and Langohr 2008: 474). We do not need to believe the facts of the ranking to take the ranking itself as a fact that exists on the social landscape.[11] However untrustworthy, as a common reference the ranking is a point of orientation from which to navigate an otherwise uncertain decision space.[12]

We do not claim that rankings solve the challenges of making decisions in a world of second-order observations. Our task is to understand *the success of rankings as a social form*, not to assess whether rankings help users succeed in making better decisions. A ranking is successful if it is taken as a reference for decisions, which can themselves be successful or not. In light of our general argument, we now turn to the 4

---

[11] The para-objectivity of ratings (the reason why they are treated "as if they were facts") relies basically on this public character: they hold as a reference as far and as long as others take them as a reference (trusting them or not) (Power 1997: 175). There is no need of objectivity to carry out this function: you just need a reference that holds as "common standard" (Langohr and Langohr 2008: 90) and is recognized as such. "What is central to the status and consequentiality of rating agencies is what people believe about them, and they act on collectively – even if those beliefs are clearly false" (Sinclair 2010: 5).

[12] Highly visible, Robert Parker's wine scores are an essential reference for everyone in the world of Bordeaux wines, even for those who do not agree with his judgment but must decide the price of the wine or plan purchases for the following year (Chauvin 2014).

dimensions of the criticism of ratings from the perspective of second-order observation. From this point of view, that they are unreliable as descriptions of independent objects is irrelevant. Their function is different and should be analyzed from another perspective, which would also be the starting point for incisive critique. As we shall argue, the same factors for which ratings are criticized appear, from this perspective, as the determinants of their problematic effectiveness.

1. Ratings and rankings simplify

Recall that critics berate rankings as quantified and one-dimensional, expressing evaluation in the form of a number that has a value only inside this ordering, lower or higher than those above or below. Such quantification simplifies and decontextualizes information, erasing all qualitative nuances.

In a complex world of second-order observations, however, this brutal simplification is necessary to bring information to a form in which it can offer guidance. That numbers are decontextualised and depersonalized has the big advantage that they are easy to export and to make public (Espeland and Sauder 2007: 16ff.; 2016: 11). They work for everyone, with no need to know where the receiver is located and what she thinks (Jeacle and Carter 2001: 300; Chauvin 2014: 33). Everyone will be free to recontextualize them as one prefers, and their ordering value does not get lost. Numerical rank takes advantage of the fact that we tend to assume that the meaning of numbers is universal and stable, independent from interpretation. Thereby, it holds for all observers regardless of their perspective, with the additional implication that, observing the ordering, you know what the others can know, even if you don't know what they think or how they reason. In the recontextualization everyone will then produce his own meanings, but you don't need to know them for the ordering to work. And there is no need to reconstruct the perspective of the issuer in order to use the ordering. Therefore the organization in ordered lists is easy to write and easy to read.[13] The users read the ordering as they want and use it as they want, stopping when it suits them and freely building their own interpretations.

---

[13] Poole (2013) attributes to this ease the rampant success of listicles, web articles presented in the form of a list.

Quantification is also criticized because it does not refer to the quality of the objects – but this again has advantages: it makes quantification compatible with irreducible singularity of the listed items. As Simmel (2004: 221) observed, money does the same, expressing the value of each object with a number (its price) and making it comparable with every other object, regardless of its quality, its emotional value, and its position on the market. The price does not take into account the fact that the money was won in gambling or was a gift from the grandparents, that the object is mass-produced or handmade, or that it can be bought on the market or, as the family jewels, is not for sale. Numbers unite and separate: each object expressed in numerical form is comparable to any other, but it is also different (bigger or smaller) from any other object.

The ordered list is a highly fungible and flexible form of organization. Rankings work because they can compare their objects without making them strictly commensurable (Stark 2011: 321), i.e., without claiming to coordinate the perspectives of the observers, which remain inaccessible.

2. Ratings and rankings are intransparent

Ratings are black boxes: their structure, criteria, algorithms, and procedures are confidential, hidden to the view and the intervention of the public. In all fields in which they operate, the power of ratings grows together with their obscurity. This intransparency is suspect because it suggests that something shady is intentionally concealed. But would genuine transparency be possible? And would it be desirable?

In a world of second-order observation, an increase in transparency is often no advantage. In auditing, efficient evaluation requires a "substantial darkness" and a certain degree of imprecision about the purposes and meanings of the enterprise (Power 1997: 16). The decisions underlying the evaluation could always be different. Transparency highlights first of all this contingency, and in many cases the spread of information undermines the trust necessary for the functioning of rating organizations, thereby increasing conflicts. More transparency does not necessarily lead to better decisions; instead, it generates a situation where informed observers can lose trust in the decision makers, who then have less authority to make good decisions (Tsoukas 1997: 840; Strathern 2000). The spread of information, moreover, has self-referential effects

that produce further complexity. As shown in the discussion about reactivity (Espeland and Sauder 2016), the rated entities observe the rating and its criteria and change accordingly - becoming more and more intransparent.

On the other hand, in a complex world of observers who observe each other and know that they are being observed, also the dependence of the evaluators from the evaluated (producing the widespread mistrust in judges who are paid by the judged) has advantages.[14] If genuine transparency is neither possible nor desirable, the form of transparency offered by ratings can become an element of self-branding of an institution. Requiring to be evaluated and willing to pay for it, an institution shows that it accepts the principle of transparency ,and this very willingness becomes an element of its reputation (Cooley and Snyder 2015). A university exposing itself to the ranking fosters a specific connotation of its image, as a college open to competition and observable in this way. You can do otherwise, but you must find an alternative way to manage the demand for transparency, and it is not easy.

3. Ratings and rankings fail as forecast

Ratings do not accurately describe the future. They often fail their predictions, lag behind, and are not able to anticipate rare events and surprises. The unreliability of ratings, however, is the flip side of their effectiveness: they cannot predict the future because, paradoxically, they actively contribute to shaping it.

Fallibility is inevitable for all measures expected to provide an orientation with respect to the social future, which should protect from surprises (Merton 1936). Ratings should facilitate the choice of a University, referring to the state of the world and to job opportunities 3 or 5 years after the date of the decision. But it is meaningless to expect that ratings, like any other measure, can describe our social future, because that future does not exist yet and will be produced also as a result of our present action, oriented by ratings. If everyone follows the ratings and chooses the study promising better chances of finding a job, at the end of the course the market will be saturated.

---

[14] A student, paying the tuition fees of his university, indirectly pays to be evaluated and to get a degree that will become part of his public image - but this does not usually leads to question the legitimacy of the teachers to grade him.

The future following our actions cannon be known in advance, and happens with or without our predictions; but if we did nothing to anticipate or produce it, surprises would simply be puzzles. In conditions of absolute uncertainty one cannot decide. But since the observers are part of the world they describe, the uncertainty they must manage is not absolute. They operate under conditions of "bounded uncertainty" (Shackle 1990: 28-48), facing a future that is unknown but bound by present decisions, a future open but not random. Of the open future we can know that it depends on what we do, even if we do not know how.

Ratings deal with this uncertainty. A decision is motivated not if it correctly anticipates the future (this can be the case of not) but if it is guided by recognizable criteria that make it not arbitrary. The future depends on our decisions, even when it deviates from our expectations and appears as a surprise. The future is not bound by our expectations, but if we didn't have them or had other expectations it would come about differently. The evolution of markets often does not match the predictions of financial ratings, but this happens because ratings have been produced and circulated, changing the structure of markets. Different ratings would have produced different markets, not necessarily confirming the ratings.[15]

Ratings offer a compass not because they make it possible to know how things will be but because they offer a standpoint so one will not drift aimlessly. If you have a financial model, for example, you can observe how markets deviated from it and learn from experience. By giving shape to our expectations about the future, ratings provide a yardstick against which we can adjust our expectations.

4. Ratings and rankings are not objective

Ratings are criticized for not being neutral observations. They describe the world from their perspective, not an independent world. Ratings offer opinions among the

---

[15] Under the label of *performativity*, such circular effects are the object of intensive research in economics: cf. MacKenzie 2006; at al. 2007; Esposito 2013. The much discussed reactivity of ratings is itself a case of performativity, which is linked to the basic condition that observers belong to the world they observe. Financial models are performative because they do not observe the world from outside. As Austin first highlighted in "How to do things with words" (1962) for the use of language, performativity arises because utterances do not only represent the world, they act on the world. Performative utterances are facts and have consequences.

others, not facts, and therefore do not seem able to provide the desired independent reference. Opinions are always contingent, i.e., they could be different. How can you rely on them and exclude arbitrariness?

But the contingency of ratings is actually a condition for their efficient functioning. Ratings do not describe; they (performatively) intervene, and thereby bind themselves. Rating practices realize a sort of neutralization of subjective elements - not a cancellation, since the judgments remain subjective, but a condition which effectively excludes arbitrariness. Lamont's study of scientific reviews, for example, shows that evaluators almost always perceive the procedures as fair, selecting the best candidates – not in spite of subjective elements, but precisely because of them: "Evaluation would be impossible without extra-cognitive elements" (Lamont 2009: 19; also Kieser 2010). Once initiated, the process binds itself and gains the necessary legitimacy that does not derive from the objectivity of the judgment (which is always debatable) but from the process itself.

Kreiner (2012) shows that the criteria that make the evaluations in architectural competitions binding and not arbitrary arise during the proceedings and cannot be established in advance. They cannot be objective and univocal, because this would unduly bind the development of the procedure. Thus, the criteria stated at the beginning will be empty and generic. The competition brief that initiates the procedure is utterly vague in order to allow the competitors to interpret it each in their own way as they develop their projects. In the course of the evaluation a project emerges as the winner. And it is this project that becomes the point of reference for clarifying the winning criteria, thereby retrospectively solving the ambiguity of the starting point. The choice at this point become motivated and almost necessary: the winning project is the one that best meets the requirements of the call, which are now clear and defined.[16]

---

[16] Schulz et al. 2001: 37 attribute to a similar effect the stickiness of rankings. Under conditions of high complexity it is difficult to understand how to deal with the heterogeneity of individual cases, and the items at the top become the reference for setting the criteria of judgment - and then they tend to remain at the top also when the criteria change. In university rankings, therefore, Harvard, Princeton, Stanford and Yale are always in the top five because they set the standard the ranking refers to.

In an enormously complex world, with different contexts and different perspectives, objectivity becomes impossible and would not even be convenient. If objectivity requires being the same to all observers at all times and in all circumstances, the claim of objectivity would become an unbearable burden, since circumstances and moments are always different and essentially unpredictable. Contingency means the ability to adapt to the context - but does not mean lack of control. Ratings and raters follow shared and controlled procedures, which make the result, although not predictable a priori, anything but arbitrary. It could have been different, but given the procedure and the specific circumstances, it emerges as correct for all participants and can no longer be changed. This allows ratings to perform the task for which objectivity is usually required: to serve as a common reference for all observers (but with much higher flexibility).

CONCLUSION: FROM THE RANKED SOCIETY TO THE SOCIETY OF RANKINGS ?

The processions of medieval Europe enacted the ranked orders of feudal society. Within each order, feudal society paraded in ranks. Among the clergy: cardinals, archbishops, bishops, priests; among the nobility: lords, barons, knights, squires; among the laborers: freemen, yeomen, servants, serfs (in the rural areas) and the ranked order of crafts (in the towns) (Duby 1980: 74, 171: Hilton 1992). Today's processions scroll past on screens: Google page ranks, Top 10 and Top 100 Lists, and the other ratings and rankings that we have argued are the guides to navigate the uncertainties of our relationship with the world, with society, and with ourselves. The two types of processions, of course, could not be more different. But yet they bear a similarity because each type of ordering provides answers (in an entirely different modality) to the fundamental questions of the order of society and of one's standing.

The social hierarchies of ancient and medieval societies, ranked from the monarch to nobles to the lower classes down to the serfs and slaves, appeared as the only possible form of order, corresponding to the general order of the universe from God to the angels to men to the animals down to inanimate objects like stones (Duby 1980: 110-19; Meier 1980). In this sense hierarchy was conceived as an objective and indisputable disposition, assigning to each person one and only one position with corresponding moral qualities and social expectations (Dumont 1980). As in the cosmos,

where origin determined the nature of things, so in social life birth determined the characteristics and opportunities of people. This disposition could not change, no more than the member of a species (a horse) could move to a different species (a dog). Individuals could have more or less different behaviors, but in this view a plebeian, as virtuous and successful and possibly rich, could nonetheless not become a nobleman no more than a decayed nobleman could become a peasant (Luhmann 1997: 689). Religious and other processions were a ritual and literal confirmation of one's standing within such a hierarchical order.

Where the social order corresponded to the order of the world, already established and known by all, rank order was the premise of observation, not its result. There were no rankings in a ranked society. The orderings of today, expressed in the rankings of evaluative lists, correspond to a very different need. Modern society is complex and contingent primarily because it does not rely on the assumption of a single order, given and shared by all. It is a heterarchical and polycentric society, in which different hierarchies and orderings intertwine and reproduce, none of which can claim to be dominant, or even to be fixed (Luhmann 1997: §§ 4.8-4.11; Stark 2009: 204-212).

Rankings, like the institutions studied in François et at. (2014), are not stable and unquestionable as the traditional rank hierarchy, but are born and die. Not all rankings are successful, and even those which do can be replaced by other guidelines. The success and ability of rankings to act as a reference do not depend as the order of pre-modern societies on the claim to correctly describe the world. Instead, they depend on the ability to build up a reference audience. One uses them "if one is not alone in doing so and can realize it" (François et al. 2014: 63).

The conditions of success of ratings are described in Lucien Karpik's study (2000) on the Michelin "red guide", which traces the history of the famous vademecum reconstructing the transformations of its function. At its beginning in 1900 it was a "technical guide" offering practical objective information for the journey, such as the availability of repair shops and service stations, the list of doctors, hotels, sites of cultural interest. The real success of the guide, however, came a few decades later, when since 1933 the guide provided an evaluation and a hierarchy of restaurants, making choices

that could be different and can be criticized (also because in the meantime competitors as the Guide Hachette began to appear). In the following years the guide became more and more contingent and subjective, but its success continued to grow. Karpik explains it by referring to the parallel development in France of a public of cultured travelers and gourmeurs, who are trained on the guide and constitute its audience. This public is not looking for objective information. In consulting the Michelin guide, its readers recognize themselves as belonging to the elite of connoisseurs of cuisine, who share the same taste, interests and criteria (labled a "likeminded community" by Jeacle and Carter 2001: 300). In reading the reviews of restaurants, the users observe and know the other readers, and eventually themselves as belonging to this group.

This kind of mechanism is not limited to gastronomic guides. The authority and authoritativeness of rankings and guides increasingly relies on participation and not on external factors (Shay and Pinch 2006). The success of the evaluation process involves the contemporary creation of an audience and of the world to be assessed, and in this view is by no means arbitrary. The judgment is correct because it provides an adequate and reliable description - not of an objective world shared by everyone, but of the specific world created by the rating and observed by its observers. When a rating is successful, its self-fulfilling prophecies become correct without being "true."[17] Who reads the Michelin guide does not only read what the reviewer who visited a particular restaurant liked or not liked, but what the whole community of readers of the Michelin guide (to which she himself belongs) knows about that restaurant and the general opinion – she is informed about the readers, not about the reviewer or about the restaurant.

Other ratings fail, not because they were not "real" or not "correct," but because they failed to attract the attention of observers and to create their own reference world. Their arbitrariness becomes then immediately apparent: why should I care for the whimsical idiosyncratic opinions of a cookery enthusiast rather than another? The same happens when an institution "dies": when a reference is no longer able to create its own

---

[17] This new field "is neither true nor false; rather, what we call 'truth' is established within this new field"(Kornberger and Carter 2010: 338).

public, one goes back to the "verdict of experience" (Francois et al. 2014: 232), that is, to first-order observation. Observation returns to refer directly to objects.

But first-order observation is not enough. We live in a world of situated and provisional orders, that hold and are not arbitrary precisely because of their contingency. This contingency also affects the structuring of individual identity. In society and in a world that do not have an indisputable and permanent order, also knowing who one is and how and where one stands is today perplexing. Ratings and rankings are tools to get an orientation in such a world – at a general and at a personal level. Whereas the ratings of Standard and Poors or the college rankings of *U.S. News and World Report* are used to observe the world that others observe, the numerous Top 10 or Top 100 lists in cultural fields are used for another need. Here the observation of others provides a reference point to observe oneself.

Consider the choice of a novel. Even when we look for a book for ourselves, as an experience or as entertainment, we cannot do so without referring to others. But we are not interested in just any others. Of interest is a specific portion of the public, namely, we want to know *what appeals to people like ourselves* – using whatever criteria we recognize and refer to when building our identity. Therefore in the cultural and "experiential" fields there are a multiplicity of different ratings, referring to different portions of the public: the ratings of the *New York Review of Books* and of the *London Review of Books* but also those of the *Jewish Review of Books*, the *Catholic Review of Books,* Oprah's Book Club list, and yet ever more finely grained niche listings. These ratings give an orientation not so much to know the world, as to know oneself through the experience of the world: a given novel, but also tasting certain foods or wines (Hennion 2015) serves at the same time to mark the membership in a group and to form one's own identity through this membership.

As was the case for pre-modern rank orders, so the order provided by rankings seems to become an indispensable reference, even if today's reference point is always changing, must continually be updated, and contributes to rather than eliminates uncertainty and anxiety. As we saw, ratings and rankings base their credibility on their ability to manage and use contingency, thereby orienting second-order observation. The

ordering they produce is not (or should not be) arbitrary, even if it depends on circumstances and is generated together with them. The order of the ranking helps then to organize the world, the relationships with others and with things. It also helps to observe oneself. While the pre-modern person always knew who he/she was and what his/her place was, privileged or not, in our society identity is increasingly mobile and negotiated, a source of ambitions and frustrations. And above all, it must always be confirmed.

Gary Shteyngart's novel, *Super Sad True Love Story* (2011) radicalizes this condition describing a world in which everyone is wearing a device (an äppärät) continuously producing a rating that can also be read by others. In this society of rankings, the rating is crucial for constructing a reference point, revealed when the crisis that closes the novel renders the äppärät useless. Without the guidance provided by the rankings, young people are caught in the grip of a modern form of deep anomie, some even committing suicide. As Shteyngart's narrator, Louie, observed about one of these young suicides: "One wrote, quite eloquently, about how he 'reached out to life,' but found there only 'walls and thoughts and faces,' which weren't enough. He needed to be ranked, to know his place in this world" (Shteyngart 2011: 270). The world "out there" offers objects (walls), discourses (thoughts), and relationships with others (faces), which are not enough as long as they remain references for first-order observation. Only at the second-order do they become significant for building identity, and for this we now apparently need ratings and rankings.

REFERENCES

Ayeh, Julian K., Au, Norman and Law, Rob (2013) Do We Believe in TripAdvisor? Examining Credibility Perceptions and Online Travelers' Attitude toward Using User-Generated Content. *Journal of Travel Research* 52(43): 437-452.

Aust, Jérôme and Musselin, Christine (2014) *The reconfiguration of the French university landscape as an indirect consequence of the Shanghai rankings, or how rankings indirectly affect the design of higher education systems*. Paris, manuscript.

Austin, John L. (1962) *How to do things with words*. London: Oxford University Press.

Beck, Ulrich (1986) *Die Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Frankfurt a.M.: Suhrkamp.

Blank, Grant (2007) *Critics, Ratings and Society. The Sociology of Reviews*. Lanham (Maryland): Rowman & Littelefield.

Chauvin, Pierre-Marie (2014) Extension du domaine de la note. Pp.70-97 in François, Pierre *Vie et mort des institutions marchandes*. Paris: Presses de Sciences Po.

Cooley, Alexander and Snyder, Jack (eds) (2015) *Ranking the World. Grading States as a Tool of Global Governance*. Cambridge (Mass.): Cambridge University Press.

Duby, Georges (1980) *The Three Orders : Feudal Society Imagined*, translated by Arthur Goldhammer. Chicago: University of Chicago Press.

Dumont, Louis (1980) *Homo Hierarchicus: The Caste System and Its Implications*. Chicago: University of Chicago Press.

Eco, Umberto (2009) *Vertigine della lista*. Milano: Bompiani.

Espeland, Wendy N., and Sauder, Michael (2007) Rankings and Reactivity. How Public Measures Recreate Social Worlds. *AJS* 113(1): 1-40.

Espeland, Wendy N., and Sauder, Michael (2016) *Engines of Anxiety. Academic Rankings, Reputation, and Accountability*. New York: Russel Sage Foundation.

Espeland, Wendy N., and Stevens, Mitchell L. (2008) A Sociology of Quantification. *European Journal of Sociology* 49(03): 401-436.

Esposito, Elena (2013) Economic Circularities and Second-Order Observation: The Reality of Ratings. *Sociologica* 2. doi: 10.2383/74851.

Esposito, Elena (2017) Organizing without Understanding? Lists in Ancient and in Digital Cultures. *Zeitschrift fur Literaturwissenschaft und Linguistik* 47(3): 351-359.

Felton, James et al (2008) Attractiveness, easiness and other issues: student evaluations of professors on Ratemyprofessors.com. *Assessment & Evaluation in Higher Education* 33(1): 45-61.

von Foerster, Heinz (1981) *Observing Systems*. Seaside (Cal.): Intersystems Publications.

François, Pierre (ed) (2014) *Vie et mort des institutions marchandes*. Paris: Presses de Sciences Po.

Grewal, Rajdeep., Dearden, James A. and Lilien, Gary L. (2008) The University Rankings Game: Modeling the Competition among Universities for Ranking. *The American Statistician* 62(3).

Gerlitz, Carolin, and Lury, Celia (2014) Social media and self-evaluating assemblages: On numbers, orderings and values. *Distinktion: Scandinavian Journal of Social Theory* 15(2): 174-188.

Guyer, Jane I. (2010) The eruption of tradition? On ordinality and calculation. *Anthropological Theory* 10: 123-131.

Hazelkorn, Ellen (2011) *Rankings and the Reshaping of Higher Education: The Battle for World-Class Excellence*. New York: Palgrave-Macmillan.

Hennion, Antoine (2015) Paying Attention: What is Tasting Wine about? Pp. 37-56 in Ariane Antal, Michael Hutter and David Stark (eds) *Moments of Valuation: Exploring Sites of Dissonance*. Oxford: Oxford University Press.

Hilton, Rodney H. (1992) *English and French Towns in Feudal Society: A Comparative Study*. Cambridge: Cambridge University Press.

Jeacle, Ingrid and Carter, Chris (2001) In TripAdvisor we trust: Rankings, calculative regimes and abstract systems. *Accounting, Organizations and Society* 36: 293-309.

Karpik, Lucien (2000) La Guide rouge Michelin. *Sociologie du Travail* 42(3): 369-389.

Karpik, Lucien (2007) *L'économie des singularités*. Paris: Gallimard.

Kornberger, Martin and Carter, Chris (2010) Manufacturing competition: how accounting practices shape strategy making in cities. *Accounting, Auditing & Accountability Journal* 23(3): 325-349.

Kreiner, Kristian (2011) Organizational Decision. Mechanisms in an Architectural Competition. Pp. 399-429 in Lomi, A. and Harrison, J. R. (eds) *The Garbage Can Model of Organizational Choice: Looking Forward at Forty*. Bingley: Emerald.

Kugel, Seth (2018, August 11) Don't Let TripAdvisor Kill Adventure. *The New York Times*, NY.

Lamont, Michèle (2009) *How Professors Think. Inside the Curious World of Academic Judgment*. Cambridge (Mass.): Harvard U.P.

Langohr, Hervig M., and Langohr, Patricia T. (2008) *The Rating Agences and their Credit Ratings. What They Are, How They Work and Why They Are Relevant*. Chichester: Wiley.

Levich, Richard M., Majnoni, Giovanni, and Reinhart, Carmen M. (2002) *Ratings, Rating Agencies and the Global Financial System*. New York: Kluwer.

Luhmann, Niklas (1984) *Soziale Systeme*. Frankfurt a.M.: Suhrkamp.

Luhmann, Niklas (1991) *Soziologie des Risikos*. Berlin and New York: de Gruyter.

Luhmann, Niklas (1997) *Die Gesellschaft der Gesellschaft*. Frankfurt a.M.: Suhrkamp.

MacKenzie, Donald (2006) A*n Engine, Not a Camera. How Financial Models Shape Markets*. Cambridge, MA: MIT Press.

MacKenzie, Donald, Mouniesa, Fabian, and Siu, Lucia (eds). (2007) *Do Economists make Markets?* Princeton: Princeton University Press.

Mennicken, Andrea (2016) Numbers and Lists: Ratings and Rankings in Healthcare and the Correctional Services. London: LSE manuscript.

Merton, Robert K. (1936) The Unanticipated Consequences of Purposive Social Action. *American Sociological Review* 1: 894-96 and 898-904.

Miller, Peter (2001) Governing by Numbers: Why Calculation Practices Matter. *Social Research* 68: 379-396.

Morin, Edgar (1985) Le vie della complessità. Pp.49-60 in Bocchi G. and Ceruti M. (eds), *La sfida della complessità*. Milano: Feltrinelli.

Musselin, Christine (2010) Universities and pricing on higher education markets. Pp. 75-90 in Mattheou D. (ed) *Changing Educational Landscapes. Educational Policies, Schooling Systems and Higher Education - A Comparative Perspective*. Berlin—New York: Springer.

Partnoy, Frank (2002) The Paradox of Credit Ratings. Pp.65-84 in Levich R.M., Reinhart C., and Majnoni G. (eds), *Ratings, Rating Agencies, and the Global Financial* System. New York: Kluwer.

Policano, Andrew J. (2005) What Price Rankings? *BizEd* September-October: 26-32.

Pollock, Neil l, and D'Adderio, Luciana (2012) Give me a two-by-two matrix and I will create the market: rankings, graphic visualisations and sociomateriality. *Accounting, Organizations and Society* 37: 565-586.

Poole, Steven (2013, November 12) Top nine things you need to know about 'listicles'. Steven Poole on the crucial facts about the internet phenomenon of written lists. *The Guardian*, London.

Poon, Martha (2007) Scorecards as devices for consumer credit: the case of Fair, Isaac & Company Incorporated. *The sociological review*, 55(s2): 284-306.

Power, Michael (1997) *The Audit Society. Rituals of Verification*. New York: Oxford U.P.

Reinhart, Carmen (2002) Sovereign Credit Ratings Before and After Financial Crises. Pp. 251-268 in Levich R.M., Reinhart C., and Majnoni G. (eds), *Ratings, Rating Agencies, and the Global Financial* System. New York: Kluwer.

Schultz, Majken, Mouritsen, Jan and Gabrielsen, Gorm (2001) Sticky Reputation: Analyzing a Ranking System. *Corporate Reputation Review* 4: 24-41.

Scott, Susan G. and Orlikowski, Wanda (2012) Reconfiguring relations of accountability: Materialization of social media in the travel sector. *Accounting, Organizations and Society* 37: 26-40.

Shackle, George L. S. (1990) *Time, Expectations and Uncertainty in Economics*. Aldershot (England): Elgar.

Shay, David, and Pinch, Trevor (2006) Six degrees of reputation: The use and abuse of online review and recommendation systems. *First Monday* 11(3).

Shear, Michael D. (2014 May 25) The Colleges Rattled as Obama Seeks Rating System, *The New York Times*, NY.

Shteyngart, Gary (2011) *Super Sad True Love Story*. New York: Random House.

Simmel, Georg (2004) *The Philosophy of Money*. London: Routledge.

Sinclair, Timothy J. (2010) Credit Rating Agencies and the Global Financial Crisis. *Economic Sociology Newsletter* 12: 4-9.

Stark, David (2009) *The Sense of Dissonance: Accounts of Worth in Economic Life.* Princeton and Oxford: Princeton University Press.

Stark, David (2011) What's Valuable? Pp.319-338 in Beckert J. and Aspers P. (eds) *The Worth of Goods. Valuation & Pricing in the Economy*. Oxford: Oxford University Press.

Strathern, Marilyn (2000) The Tyranny of Transparency. *British Educational Research Journal* 26: 309-321.

Stuart, Debra L. (1995) Reputational rankings: Background and development. *New Directions for Institutional Research* 88.

Tsoukas, Haridimos (1997) The Tyranny of Light. The temptations and the paradoxes of the information society. *Futures* 29: 827-843.

Elena Esposito is Professor of Sociology at the University Bielefeld and at the University of Bologna. Her published work engages topics such as the theory of social systems, media theory, memory theory and the sociology of financial markets. Her current research addresses a sociology of algorithms.

David Stark is Arthur Lehman Professor of Sociology at Columbia University where he directs the Center on Organizational Innovation. He is also Professor of Social Science at the University of Warwick. His current research project, "Diversity and Performance: Networks of Cognition in Markets and Teams" is supported by a five-year, Advanced Career Award from the European Research Council.