

Article

# How and Why the Metric Management Model Is Unsustainable: The Case of Spanish Universities from 2005 to 2020

Lucas Jódar <sup>1</sup> and Elena De la Poza <sup>2,\*</sup> 

<sup>1</sup> Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, 46022 València, Spain; ljodar@imm.upv.es

<sup>2</sup> Centro de Ingeniería Económica, Universitat Politècnica de València, 46022 Valencia, Spain

\* Correspondence: elpopla@esp.upv.es

Received: 24 March 2020; Accepted: 23 May 2020; Published: 28 July 2020



**Abstract:** The metric management model is a method based on quantitative indicators called metrics and is used to evaluate individuals and organizations. Organizations' sustainability is related to risk and expectation concepts and both are, in turn, related to the metric management model (MMM). The main objective of the present research work is to analyze the MMM applied to the Spanish university system (SUS) and the propagation of its consequences. The secondary objective is to study alternatives to the metric management system applied to the SUS to avoid its negative socio-economic consequences. Our results reveal how applying the MMM to the SUS, based on the metric evaluation and the ranking monitor model, deteriorates research quality, students' levels of education and working people's well-being at university. Finally, university managerial boards, teased with the "mirror" of university rankings and the picture a simulacrum of reality, are still unaware of the damage.

**Keywords:** sustainability; metric management model; university ranking; labor welfare; patriotism; public management

## 1. Introduction

Sustainability is, to a certain extent, a synonym of the expression "resistant to the pass of time", and is durable, but also implies the environmental conditions of an area that allow the population to live in a clean and healthy environment during a given period of time because, otherwise, we would be talking of its abstract meaning. We think that sustainability is not only related to the physical sphere of its meaning, but also to the mental state of the involved population. We claim paying attention to intangibles like labor environment, labor stress, motivation or excitement. As countries do not disappear, a country's unsustainability means its irrelevance.

We ought not to forget that humans are driven by emotions and we are social beings who imitate our neighbors and colleagues [1–3]. It is clear that sustainability means that deterioration must be limited or controlled, or even driven out. Sustainability is also linked to prevention because potential damage should be identified in advance.

Sustainability is also related to risk and expectation concepts, both of which are related to data, quantification, numbers and probabilities. Numbers act like a knife that can help to cut things, but we can cut ourselves when we use it. The main risk of numbers is their use and interpretation.

For the sake of clarity, it is worth pointing out some of the concepts used throughout this paper. A model is an approximation of reality. Management is a set of principles related to the functions of planning, organizing, leading and controlling, and the application of these principles in harnessing

physical, financial, human and informational resources efficiently and effectively is done to achieve organizational goals [4].

Management is essential for an organized life and is necessary to run all types of organizations. Managing an organization means getting things done with and through other people to fulfill its objectives. The management of some individual activity needs to be evaluated according to suitable indicators that measure the accomplishment of targeted goals. Managers are individuals in charge of this activity.

The metric management model (henceforth referred to as MMM) is a method used to evaluate individuals and organizations and is based on quantitative indicators called metrics, which summarize right behaviors with a numerical score. The management of a public organization is very different from that of a private one because the sustainability (survival on the market) of a private organization depends on profits, while a public organization is subsidized by public funding (taxpayers or public debt). Consequently, unsustainable private organizations disappear, but public ones do not, while the support of public funding continues [5].

It would be too arduous and time-consuming to define the targets of a public institution like a public university [6] but, in some way, these objectives should be related to the search for excellence in research and education to improve a country [7–9]. This close relation between a university's performance and a country's behavior (economic and socio-political) is discussed by Hedges [10] and Denault [11], but the latter goes further into this rationale by linking the growth of mediocrity with university performance.

Public funding involves searching for social return because public money comes from the taxes charged to citizens. The social return level also depends on the type of government, time and location. This paper focuses on the Spanish university system (SUS) during the 2005–2020 period, [12]. Most of the analysis and conclusions can be extrapolated to many countries, mainly those in the western orbit.

In theory, management of public institutions is controlled by the accomplishment of the strategic planning devised by governments and public authorities. Yet in practice, the accountability of managing a public organization is fuzzy because stakeholders who claim the responsibility of optimizing taxpayer contributions in socio-economic return terms are lacking [13]. The political nature of public management boards explains this situation, which worsens due to not only the instability of governments and budgets, but also to slow decision-making processes. Thus, public university managers, who are surprisingly elected as professional politicians [13], play the role of amateurs who like to be professional, and not professional managers, but professional politicians. So, expectations of public university management being right is as naive as expecting a country's management to be right, and both are chosen by criteria of expertise regarding how to successfully run organizations.

As with humans, at least in western countries, once a need is not met in a free market economy, an opportunity of profit appears and a new business emerges. A lack of clarity in defining a public organization's objectives favors new business growth, which consists of "private" agencies qualifying organizations, measuring university performance and devising rankings of universities [14]. Even when controlling university performance, it is useful for understanding an organization's processes and for allocating resources with the continuous improvement spirit. However, it is not used only for internal purposes, but to also compare institutions with contemporaries. Hence, these rankings qualify all public and private worldwide universities, regardless of their legal form, size, geographic location, and the origin and volume of their funding. In addition, cross-country differences between university systems are a fact [15,16]. So, it is neither easy nor feasible to compare higher education institutions (HEIs) with rankings [17].

The dysfunctions of rankings that qualify both public and private universities, and the ambition of amateur politicians, motivate their excess monitoring of university management according to rankings [18,19]. Managers of public universities, with the help of big data and artificial intelligence, have found a new role to play—the completion to improve rankings [13,20]. This is an exhibition of good public managers and as a result of managers promoting their own images.

A private higher education institution (HEI) will act as a provider of higher education (HE) financed by public funding, but will also provide HE to those individuals who can pay the cost of its services [21].

Once again, the parity risk/opportunity emerges. The risk of the unsustainability of the public university system coexists with the opportunity of private ones, mainly if private ones start attracting some of the best professors of public universities with good contracts, or if they take advantage of the nonsense compulsory retirement of the public university to attract them [22,23].

Numbers are seductive, look objective, are transparent, summarize information, look scientific, and seem almost perfect. However, this is not the case, and can even be dangerous if misused to summarize complex phenomena related to human behavior—see [6,10,24–38].

Critics who use numbers state that they avoid profound issues like considering intangibles in favor of metric indicators that are more objective; see page seven of [26]. The use of metric management suggests a preference for objectivity as opposed to depth. As [39] and [40] discuss about the brain's uncertainty of knowledge, this comes over increasingly more evidently in the simulacrum era [26], where lies the main strength that drives the world [20]. Apart from the theoretical point of view, the practical implementation of metric management has important dysfunctions that discredit the approach.

In fact, using the MMM becomes a pay-for-performance method to measure human activity, and this awakens lots of human passions by distorting the procedure. The individual application of the MMM allows promotion through a candidate's positive accreditation and, hence, the procedure becomes a pay-for-performance method. The application of this model to public organizations is very risky and complex, and here we explain why it is wrongly managed.

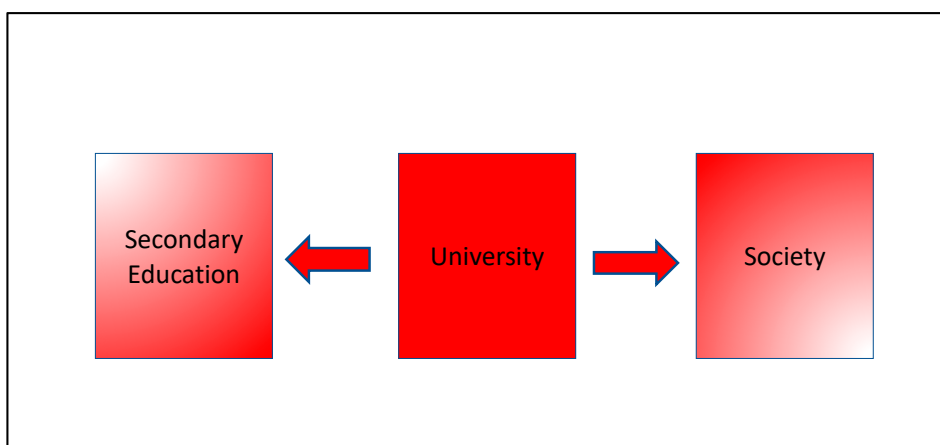
Numbers and algorithms can replace neither profound knowledge and understanding [26], p. XV, nor individuals or organizations. At the most, this information can be complementary to, but not essential in, the study of human behavior.

The main objective of the present research work is to analyze the MMM applied to the SUS and the propagation of its consequences. The secondary objective is to study alternatives to the metric management system applied to the SUS to avoid its negative socio-economic consequences. This paper should be regarded as a preventive study against the growing deterioration of the SUS, which has taken place since approximately 2005. Although the main driver of this deterioration is the mimetic behavior described by the McNamara model 40 years ago in the USA, the new business of big data and university rankings [14] contributes to the public university's deterioration, with the corresponding degradation of the education and training of most of the country's elites. This phenomenon predicts a continuous wave of mediocrity that spreads outside the academic world to companies, civil society, etc. [10,11].

In addition, the wave of mediocrity that starts from the university floods the previous levels of the education system through the education and training of the graduates in charge of the levels of education that come before university learning.

This phenomenon is like a contagious infestation of mediocrity that spreads to the country's whole fabric. In fact, the matter of MMM is a social illness, inspired by some sort of quantification ideology, which spreads by social contagion to society as a whole. The severity of this social illness is paramount for two reasons. The first is that social illness is not identified as a real illness. The second is that the agent vectors of this social illness are not only unaware of this, but even presume and flaunt it. The situation is like that of someone suffering a serious illness, but believes that (s)he is healthy and does not much care about his/her life.

The propagation phenomenon is like fine rain: it never stops, but ends up wetting almost everything [41,42]. Although MMM dysfunctions are spread by public universities, their effect could reach the management of other sectors, such as medical activity or police statements, and would critically affect both the public health service and security if the MMM is also applied by other amateur managers who are elected in political ways (Figure 1).



**Figure 1.** The spread system of the metric management model (MMM).

This paper is arranged as follows. Section 2 explains how the MMM originated and why it was questioned in the US army. Section 3 identifies the set of main dysfunctions of this model's application and its consequences. Section 4 provides an alternative way, with different degrees of intensity, to manage and govern higher education without using the MMM fully, but only a small part of it. The paper ends with a conclusions section that summarizes the impact inside and outside the university, as well as the effects of the MMM from the mental health point of view for both teachers and students. The potential utility of the suggested changes to be made to the MMM is not only for the Spanish higher education system, but also for other countries where the model is used.

## 2. Background

### 2.1. The Origin and Rise of the MMM

It is difficult, or impossible, to identify the origin of ideas and their author or owners, but we can approximate the origin of concrete actions through history in western countries. The ultimate origin of the MMM dates from 1862, by Liberal Parliament members in Britain, who proposed a new method of government funding for schools. In 1911, American engineer Frederic Winslow Taylor coined the term “scientific management”, based on attempts made to replace the implicit knowledge of a factory's workforce by mass production methods, which were developed, planned, monitored and controlled by managers [27]. This is closely related to the industrial revolution, the division of labor and its accountability.

It is important to point out that these management ideas were implemented to evaluate the mechanical activities associated with industrial activities, where the workforce performed the same activity, done in a factory, and where creativity and imagination were scarce, or nonexistent. It is worth stopping here for a moment because one of the keys of the MMM is to compare noncomparable activities. One can measure, for instance, the performance of a workforce collecting oranges in Valencian fields, or the number of documents processed by administrators in a company. Yet it is rather dubious to export these quantitative criteria, and to compare different people's creative activity performed in distinct places under varying conditions.

The pay-for-performance criterion of evaluation and comparison can be implemented when conditions and activities are essentially the same. The activity of a teacher at school in a very poor area of a country with very limited resources and with a proportion of students who have behavioral problems and different ethnic origins and races is not comparable to the activity of another teacher from a different country in a school located in a rich area with good resources, having half the students of another teacher, and where students have no significant behavioral problems.

The essential point of the starting conditions must be identical in order to compare human activities because, otherwise, one compares what is not comparable and the results are simply lies. International indicators and rankings usually make this big mistake and the conclusions are not reliable.

The Dean of the School of Education at Stanford University reflected the need for more efficiency through standardization and monitoring following Taylor's method in his influential book [43]. Taylor's management model was increasingly adopted in many factories in the USA between 1919 and 1939, and it was the norm in companies like General Motors in 1950.

Apart from engineers, Robert McNamara extended Taylor's scientific management model to a myriad of organizations. McNamara applied the MMM during the 1941–1960 period at the Harvard Business School. McNamara taught at Harvard until 1946, when he moved to work as an executive for Ford. He later carried out the message of metrics (by 1960) with the largest organization in the USA: the US Army. From 1968 to 1980, McNamara was executive and President of the World Bank. During these four decades, American business schools were transformed. In fact, before McNamara, no consensus had been reached about which management skills should be taught. From the 1950s onward, the ideal business-person was a general manager equipped with a set of skills that were independent of concrete industries, and focused on the mastery of quantitative methodologies.

As usual, the new ideas developed in the USA by imitation and contagion spread from this country and, of course, arrived in Spain years later, around 1975, due to the scientific isolation that took place during Franco's dictatorship in Spain. Thus the late arrival of new model ideas to Spain, combined with the additional delay of the university system's transformation, meant that ideas of the model started to arrive in Spain around 1990 which, as we go on to explain, were questioned and partially rejected.

McNamara epitomized the hyper-rational executive who relies on numbers rather than intangibles like trust, resilience or emotions, and who applies this metric management to any organization.

In fact, McNamara exported Taylor's scientific management to organizations and outweighed the use of metrics from manufacturing companies to any organization: the contagion was consummated. It is worth remarking that this occurred in the USA within a time interval (1950, 1980), with acceleration in the first half and deceleration in the second half, due to an event that we now explain, which opened many people's eyes to the other side of the Atlantic.

In the USA, the intensity by which the MMM's ideas were diminishing, but continued to spread by followers and copiers, as is usual in Spain, was copied with some delay by the public sector. The Spanish private sector adjusted the errors identified in the USA by 1990 because the sustainability of the private sector means market survival. However, in the Spanish public sector, managers are not professionals, but amateurs, and the correction of mistakes arrives late, wrongly or never at all.

This is one of the main reasons to explain how an economic financial crisis affected Europe; Spain suffers consequences to a greater extent than its neighboring countries. Politicians and public managers identify errors late and, as suggested by J.F. Revel in [20], when they possess knowledge, they do not implement changes because their interest lies in the short-term and their lack of patriotic sense makes them look elsewhere. It is not merely a matter of ignorance, a large component is due to personal interest and a lack of patriotism comes into play.

## 2.2. Questioning the MMM

This next section will cover why the MMM errors in the 1970s in the USA were questioned.

As we will soon note, the MMM was criticized in the USA when Robert McNamara was nominated as the Secretary of Defense by Kennedy in 1961, the Vietnam war lasted 6 years, and Kennedy wished to get the Viet Cong to the negotiation table. As the Vietnam conflict escalated and the USA sent more troops, it became clear that this war was a war of wills and not a territorial war.

The metric management used by R. McNamara to measure the Vietnam war's progress consisted of quantifying the number of enemy soldiers killed. It is necessary to pay attention to this fact because it is crucial to select the relevant item to be measured in the MMM. Thus, it is no trivial task. This is

regardless of it being possible, or not, to select an appropriate variable to summarize the complexity of the problem.

The main problem of McNamara's approach is that the crucial variable, apart from the difficulty of measurements, is not the enemy being killed, but probably the capacity of the enemy's resilience. McNamara mistakenly selected a wrong variable, missing the question of whether or not the enemy's resilience is a measurable variable.

A war is not won until the enemy stops fighting. This also occurs in sport competition, mainly in individual sports like tennis or boxing. In collective sports, this is not so easily observed because one needs a collective coordinated mental behavior to continue fighting.

The metric variable used to measure progress was the number of enemies killed—body counts were published daily—but it was not a useful measure to win the war because Vietnam's resilience capacity was greater than the number of American victims that society was capable of affording. Apart from this, many data were fake given McNamara's incredible personal interest.

The use, abuse and misuse of data by the American military during the Vietnam war are a troubling lesson about the limitations of information as the world hurls toward the big data era. Underlying data can be of poor quality and also biased. They can be wrongly analyzed or even be used misleadingly, and can fail to capture what they purport to quantify [44].

McNamara's prestige increased his mastery of system analyses: "the business of making sense of large organizations" [45]. He thought that only the supremely rational approach using the data of every facet was relevant by finding simplicity in complexity [27,45]. This approach fails to manage human behavior, simply because humans are driven by emotions [1,2], which must be taken into account in the model.

Intangibles like greed, fear, self-esteem, confidence and human mimetics are also extremely important in the analysis of human behavior. We could state that human organizations and individuals cannot be described as Markovian processes, where measurable data at a given time collect all the relevant information.

The example of McNamara's mistake when applying the MMM to the progress made in the Vietnam war should alert the followers of big data, artificial intelligence and machine learning. Emotions must be taken into account, and one crucial question is: how are emotions measured through data? Furthermore, if important intangible factors are not measurable, what needs correcting in the model?

McNamara, who died in 2009, acknowledged his mistakes and thought that many statistics were misleading or erroneous [45]. He was so set on data, and so obsessed with the power that they offered, that he failed to appreciate their inherent ability to mislead.

Thus, the failure of the MMM developed by McNamara when applying it to measure the US Army's progress during the Vietnam war was a turning point for the model's credibility. This time coincided with the war's end in 1975.

Forty years later in Spain, managers of public institutions, like ministry agencies and universities, are managed by amateur leaders who ignore the lessons learned in the USA after McNamara's management of the US Army during the Vietnam War.

Unfortunately, McNamara's influences did not end with the wrong management of the US Army because he occupied the presidency of the World Bank in 1968 and continued to do so until 1981. Today, the World Bank is a generator of global indicators that irradiate misleading influences that should at least be reflected on [24,25,27].

### 3. Methodology

The methodology applied in this study consists of qualitative research based on a deductive approach. First, the application of the MMM to university systems is analyzed according to the existing literature to identify the core dysfunctions of the metric valuation method, which mainly focus on Spanish universities.

Once the dysfunctions of the MMM in the public university system are identified by applying Campbell's and Goodhart's laws, the next study step consists of analyzing the main drivers of the SUS to explain the damage of applying the MMM. Hence, this analysis allows searches for short-term alternatives to be implemented into the Spanish public university system.

The typical dysfunction pattern of the metric valuation of human performance was formulated in 1975 by two social scientists operating on opposite sides of the Atlantic in completely different environments. Starting with the discovery made by the American social psychologist Donald T. Campbell [28,35], the so-called "Campbell's Law" states: "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more it will be to distort and corrupt the social process it is intended to monitor". A variation of this is "Goodhart's Law", named after the British economist who formulated it in 1975 [29,30,35], which states that: "Any measure used to control human behavior is unreliable" or, in other words, "Anything that can be measured and reward to humans will be gamed"; see page 20 of [27,30,35].

Knowing that humans are moved by emotions [1,2], and that the performance of metrics involves payment for performance, in some direct or indirect way both the policy makers who ideate metrics evaluation, and the followers and copiers like ministerial agencies and rectors, research vice-rectors and involved stakeholders, cannot make a naïve interpretation of human behavior and the wrong collateral effects of metrics evaluation programs.

Both Campbell's and Goodhart's Laws came about after carefully studying outstanding people who know human weaknesses. One needs psychology, sociology knowledge and a dose of skepticism when you have no human contact when using metrics valuation models based on mechanical procedures. Below, a wide spectrum of examples is shown where the criterion is gamed, and they justify both Campbell's and Goodhart's Laws.

Of course, there are companies that conduct business with this affair by selling programs, rankings and so on, and they seek profit like any other business does [14]. However, the public managers and citizens who support them with the taxes they pay require the right accountability.

In addition, the wrong measures that place the Spanish public university system's sustainability at risk are important enough to pay attention to this matter. However, deterioration goes on in the Spanish public system. Nobody pays attention to Campbell's and Goodhart's Laws, and the public system fits the MMM in every way.

## 4. Analysis

### 4.1. Dysfunctions of the MMM and Its Effects against the Public University System's Sustainability

By extending the comments of [27], and by taking into account the particular situations of the SUS, below we list the main set of dysfunctions of the metric valuation method, which focus mostly on the university environment. Starting by the distortion of information:

(i) *Measuring the most easily measurable, easy versus complex.*

It is clear, for metric management followers, that they will not measure the intangibles that are very difficult to detect and quantify. Perhaps they will not even think about their existence at all. They are used to being fooled by rankings and indicators, and do not think about anything different to what they measure.

There is a natural tendency to simplify problems by measuring the most easily measurable elements. Yet what is easily measured is rarely what is more important, and may well not be important at all. This is the first metric dysfunction type.

In Spain, a researcher's performance, in agreement with the MMM, is measured in terms of the number of papers published in ranked journals, and the number of times an author is cited [46,47].

This is not a correct assessment indicator because, without carefully studying the published papers, it is not possible to identify the scientific contribution or the added value of a paper and its applicability

and/or effects compared to previous studies. It has also been proved that the citations a paper receives does not correspond to the number of times that a paper has actually been read [31,36,38,48,49].

The economic incentive based on pay-for-performance motivates an almost general tendency to maximize productivity by repeating an idea, in a more or less hidden way, to reach another publication that maximizes the economic return. It is a kind of publication greed [48–50]. The present Spanish university environment forces overproduction, but hides the poor quality of contributions; see page 122 of [32].

If anyone wishes to know someone's research activity, it is necessary to carefully read all his/her produced works and to identify the originality, level of redundancy (which is usually high), and also if the candidate is able to manage and solve different problems. The mechanical procedures involved in the MMM do not capture these fine intangibles.

This "supermarket" research model is what the SUS promulgates and what it should resist. The system enhances the quantity by objectifying research, knowledge and the studied subject; see page 58 of [32]. This paradigm means that researcher activity lacks interest and intellectual transcendence. As a result, researchers speed up and publish as many papers as they can [51].

Thus, the system only values research output, understood as published papers, and disregards the scientific and social interest. The consequence of this system is that many researchers exist who essentially perform the same activity throughout their scientific careers, with the inevitable component of repetition and a lack of innovation [52]. One major problem is not lack of production, but an excessively repeated and irrelevant production.

Candidates should be asked to explain the relevance of their production. No doubt many researchers who are the authors of lots of papers would probably not know what to say. The ranking of a journal does not guarantee the quality of a paper published in it. Moreover, the gap between the prestige and relevance of, and the interest in, lecturers' teaching and researching activities widens with time, and teaching is increasingly disregarded [51].

Thus, attention should focus on teaching. Basically, the teaching process is assessed through teaching publications, teaching experience (number of years teaching) and sometimes questionnaires passed to students to assess their opinion about a teacher's performance. Students' opinions matter, but it should be measured by avoiding any conflicts of interest that exists when surveys are given during the teaching period. Surveys should be given after the teaching period, so that students develop a critical perspective opinion about the lecturing they have received.

Thus, the teacher who gives students higher scores usually gets better results in assessments. The number of years teaching is no guarantee of knowing how to teach a class. An example of the teaching class needs to be evaluated with an expert committee.

A good teacher can be identified only by checking a class in vivo, and by knowing not only the technical information transmitted, but also how to introduce the subject, how to motivate and how to proceed with the teaching-learning process. These issues are not considered by the MMM.

*(ii) Measuring inputs rather than outcomes.*

As followers of the MMM need to obtain numbers, the quantification of easy indicators never ends. This even implies the confusion of outputs with outcomes. This misleading behavior is embedded in the Spanish university environment. Hence, those people who have not been qualified for promotion, have research grants and are starting out in their academic life, attribute themselves with quality and excellence. In this context, captured economic resources are the equivalent of excellence. Gaining grants is acknowledged by the university and by government agencies of evaluation.

Excellence is the habit of doing a job well and sustainably, according to Aristotle [53]. Money implies neither quality nor excellence. Another thing is rectors' love of money incomes.

This applies to not only at the individual level, but also the organizational level. Individuals argue that they are excellent because they have economic resources gained through grants and competitive projects funded by public authorities.



Nobody complains about the confiscatory behavior of the business manager of a university. Commissioned people must pay everyone from their pockets every time they travel, while rectors spend large amounts of money on nonsensical celebrations, like the university's anniversary, which is often younger than many of commissioned people are.

Why? This is not because working people at the university agree with this procedure. However, as the university only breathes the air of amateur politics, those who could complain represent the majority, but they prefer to remain silent so as not to risk a project, a position, or a medal. As I wrote in [13], the democratic election of rectors is an unfortunate error that stems from excessive democracy after Franco's dictatorship period. Rectors do not make right decisions, but convenient ones, to aid their interests and their public image, and these decisions are far removed from public or patriotic targets.

(iii) *Degrading quality through gaming and cheating.*

Quantification is seductive because it simplifies and organizes knowledge by offering numerical information that allows easy superficial comparisons to be made between individuals and organizations. In fact, quantitative indicators are used to accredit individuals and to build a ranking of universities [9]. Yet such simplification may lead to distortions as making things comparable often means they are stripped of their context, history and meaning [34].

This results in information looking more certain and credible than it actually is: caveats, particularities, ambiguities and uncertainties are peeled away, and nothing does more to create the appearance of being certain and objective than expressing information numerically [37,38]. Campbell's and Goodhart's above-cited laws warn about the inevitable attempts to game the metric when much is at stake. Gaming comes in a variety of forms.

The main focus of dysfunction is often fraud in some official authors' multiauthor academic products: paper, book, or PhD dissertation supervision. The standardization metric approach, without the candidate's personal judgment, is absolutely unable to identify this frequent dysfunction.

A naive glance at this fact may lead us to think that there is no advantage for all the remaining authors if an unrealistic author is included in the publication. In fact, this is naive because no teams operate as a clan. If some cooperators are at the top of their promotion level, they really do not lose anything at all by being a co-author with others, regardless of them being students or colleagues.

In fact, each unrealistic author improves his/her curricula, and remains faithful to some corrupt boss. In addition, unreal authors increase the clan's dependence and they allow themselves to cooperate. This gaming is systemic, not accidental, in the Spanish university, which is full of mediocre people in all categories as a result of the soft MMM operation.

A second gaming dysfunction level is the repletion of ideas in several publications in some hidden way. If a committee does not invest time in carefully reading all a candidate's publications, it is impossible to identify this repletion. Credit placed in journals is empty for several reasons. The first is that a referee's job is not done professionally at a high rate because it is hard work that goes unacknowledged and is, thus, understandably quite disregarded [49].

Another dysfunction is the market of citations; see [30,36,48,49]. As the citation index is incorporated into metric evaluations, and as Campbell's and Goodhart's laws warn, there are people who merchandize citations to improve their citation index scores. Some simply ask you to include them in the citations of publications. Someone's lack of scruples with regard to citations can also be seen when they play the referee's role, as they incorporate the obligatory inclusion of several citations of themselves into their report. I have personal experience of operating as an editor, where a referee claimed a "sine qua non" condition of acceptance; the inclusion of 13 references by himself.

Gaming examples are not exceptional, and even laureated people with honorable prizes are at risk of such practices. Some authors claim their authorship in 30, 40, or even 50 papers, in good journals, per year. In some places, bosses force all their team members to sign him/her as the author of most of the team's publications [30,38,49].

In other cases, the strategy is more equalitarian for as many members as possible to achieve the top category. For instance, in line with the clan strategy, one helps to first select someone to be promoted. Having met that target, the next person on the clan's preferences list is promoted, and so on.

The strategy is to artificially choose someone to be the co-author of all the publications, who is selected in order of preference. Thus, the MMM is not only based on personal judgment, but also comes with standardized mechanical criteria accredited to all clan members. The culture of speed and quantity boosts unethical behavior, as Campbell's and Goodhart's laws explain.

*(iv) Improving numbers by lowering standards.*

Spanish public universities are financed partly by the number of graduates in their first year. This quantitative criterion appears, in principle, appropriate insofar as the university's success should be related to its success, shown by the number of its students gaining a degree. However, this clear example of pay-for-performance suffers the effect of Campbell's and Goodhart's laws.

In fact, the indicator measuring the number of graduate students is misleading and generates a cobra effect because rectors spread the idea in all schools or faculties that the number of graduate students should be high because budgets depend on this. Deans of faculties transfer this idea to all lecturers, and if the number of graduate students is lower than expected, they are alerted and advised to improve their success rates.

As student culture is to select the minimal effort itinerary, the conclusion is that, in order to maintain the success level, one has to cut the level of requirement each year by lowering the standards to pass each subject. This conclusion means a permanent slowing down of the standards of requirements. Hence, one experiences a lowering contagion wave in agreement with both the rector and dean's wills. Implementing lower requirements is consummated yearly. Public universities are institutions that provide increasingly devaluated degrees.

Negative MMM effects do not relate only to the slowing down seen at the academic level that each experienced teacher has experienced but, more importantly, to the involved minimal effort culture that produces a fragility of character, lack of resilience to get over the natural obstacles of real life and weakness of character induced by mollycoddling education, which all contribute to increased mental troubles in young people [54]. It is worth pointing out that the high unemployment rates in Spain contribute to the labor shock because people eventually graduate, and even those who get a job are very badly paid. It would seem that the distortion of young people minds' due to the MMM effect and the minimum effort culture also occurs in the USA, as found in [44].

In addition, as the level of interplay between the university and companies lowers and is not compulsory, graduate students do not know the tasks they must perform in a job and, as a result, jobs on the market pay lower wages. Companies are not willing to pay high wages to individuals who ignore their job. We should also take into account the high number of graduate students because, in Spain, families are influenced by the Veblen effect [3] and prefer to send their children to a public, but also cheap, university. However, there being more university students in the public sector also depends on the way the economy behaves and on it being steadily maintained in the private university sector [23].

This is the consequence of a wrong financing criterion that stems from the government, and is stimulated by rectors, deans and most research and teaching staff, who are made prisoners of the model.

The MMM's dysfunctions and the pay-for-performance criteria are listed above. The teaching and research staff at a university are stressed by more and more demanding metric indicators based on the quantities of production items [46,55]. As rectors and vice-rectors of research are fooled by rankings, and are also driven by the same metric model, the demand for numbers increasingly grows. So, researchers become increasingly stressed, and teaching becomes more disregarded, because research is less important than researchers being promoted [46,56,57].

Note that very few people have the freedom to disregard all of the metric criteria because, apart from full professors, the remaining teaching and research staff need to be evaluated by the ministry agency and are, thus, forced to follow the mainstream, unless the paradigm changes.

Thus, lecturers who wish to be promoted become slaves of the metric evaluation process. The only ones who are free from this pressure are full professors. However, as they have students and cooperators who find themselves in the processes of promotion and constant accreditation, even full professors feel forced to stick to these rules.

So almost everybody is forced to follow the wave of metrics. The very few full professors who have the possibility of being free know that if they do not keep quiet, they go against the policy of both the university and government agencies. They know that they may lose a project, a position or a medal. In fact, we are witnessing a “too big to fail” situation.

We are in fact suffering a metric management epidemic whose transmission vectors are rectors and ministerial agencies, and whose victims are not only the majority of teaching and research staff, but also students [57].

The contagion extends because the younger people are, the higher the influence of the metrics is on them. The worst fact is that the youngest learn the message of quantity, speed and projects having no risk in terms of guaranteeing productivity, which is exactly the opposite of the advice given by Ramón y Cajal [50]. This kind of supermarket model based on quantities, metrics and rankings does nothing else but deteriorate both the research and education quality, and stresses everybody [58].

The nonsensical labor stress motivated by this research competition with regard to rankings and quantities discourages excellence, an effect of the quality and slowness habit [59]. The MMM promotes quantity and speed, and is most unhealthy in both physical and mental terms. Rates of excessive working, depression and separations soar.

To the stress experienced from research competition, we must cover the asphyxiating bureaucracy that is partly motivated by the New Contracts Law [60], and partly by the iron instructions given by the university’s business manager to the administration staff which, in practice, works more against the teaching and research staff than it does to serve. Actually, the rules imply that anyone commissioned to travel is owed service commissions, but in fact they are made to pay with their money. Thus, many researchers decide not to travel, and are not motivated to work or apply for projects.

The rates set per country to refund the accommodation and subsistence allowances to commissioned people incurred while traveling have not been updated for 14 years in Spain [61], and commissioned people frequently pay part of the compensation. This issue does not appear in the rankings, everybody silently complains, but nobody does so publicly. Some kind of fear stops critics because, in some cases, a complaint coexists with administrative repression.

Rectors should acknowledge in international G7 meetings that one talks of not only a country’s gross domestic product (GDP), but also the happiness index of countries. They could take this index into account at universities because the deterioration of the welfare inside universities is considerable. The situation would probably be different if lots of people complained, but fear or interest paralyzes them.

Of course, while the business manager’s rules harm commissioned people, the rector spends big money purchasing expensive rankings [14].

The university environment cuts the excitement of starting new projects short. For the reasons and effects explained above, the actual model is unsustainable. In fact, in the USA, where the model was created, it is no longer applied because it failed in managing the Vietnam War.

However, deterioration goes on in the Spanish public system. Nobody pays attention to Campbell’s and Goodhart’s laws, and the public system fits the MMM in every respect.

The question is, will the Spanish public system survive with this model, which was invented in the USA, and was rejected and has been questioned since the 1980s? We will attempt to explain this in the next section.

#### *4.2. Alternatives to the MMM in the Public Higher Education Spanish System*

The MMM is not only applicable to public higher education systems, but also spreads in most management activities, with different degrees of willingness and intensity in many countries. Its impact

is due to contagion because it is spread by organizational leaders, and is more or less imposed on organizations throughout incentive policies. A quantitative ideology exists that should not be applied loosely because it involves serious risks wherever it is implemented. For instance, the MMM applied to assess police officers' performances by measuring the quantity of discovered crimes may easily produce the consideration of innocent citizens as guilty ones. Another example of an MMM application would consist of a surgeon making evaluations based on the number of successful surgeries, which would tempt the surgeon to decline a high-risk surgery. The MMM is not a good alternative to judgment, and using metrics requires judgment about whether to measure, what is important and if it is measurable, and how to evaluate the significance of what has been measured.

The questioning of the MMM in the USA was started by the military staff of the US army [62]. Robert McNamara, the MMM intellectual creator, was removed from his position as Secretary of Defense in 1968 because of a lack of accuracy in the metric quantitative indicators used to analyze the development of the Vietnam War, which McNamara admitted in his own bibliography [45]. Then he moved to the World Bank, where he presided until 1981. During this period, he spread the ideas of the MMM around the world. This is why, albeit questioned by the military management and the Pentagon [62], the MMM was exported abroad, but was also implemented in the USA in the education and economic sectors. This is natural because of the political influence of World Bank policies. Either willingly or by contagion, and more or less consciously imposed, the hegemony of the USA spread the new methods in the western orbit of influence.

The MMM in Spain involves a complete set of nested measures from the national government, others from local governments and, finally, internal ones inside one university where the institution's presidency possesses a certain level of autonomy.

Thus, it is not a matter of accepting or rejecting the MMM. Its relevance relies on the critical issues considered and their degrees of intensity in applied metrics and measures.

This means that the MMM is not a plan strictly conceived from the top and does not have a clear origin, but is a process that has accelerated in the last two decades in Spain, and has been spread by a kind of contagion effect through the Council of Rectors of Spanish Universities (CRUE).

Unfortunately, education and research have not ever been priorities for Spanish governments since democracy arrived in 1978. Proof of this is shown by the eight different National Laws of Research and Education in the last 40 years. The two main parties that have governed the country, the left-wing party (PSOE) and right-wing party (PP), have been unable to reach an agreement to build a stable research and education law.

This is one of the reasons why the MMM is a product of the national agencies of evaluation of research and education, mainly of the CRUE. The drivers of the CRUE are much stable than those of the national government, which more or less change every 8 years with some exceptions. In fact, when the national government changes, the names and competencies of the ministry or ministries in charge of research and education also change. Rectors may survive changes in national government composition, as does the CRUE, but are presided by the Spanish Ministry of Education and Research Affairs.

The core damage of the MMM to the SUS lies in three main issues, for which short-term alternatives can be implemented. We go on to explain all of this damage and its consequences (I-III). From these three main causes, other indirect damage spreads to the overall public higher education system, including public universities. Apart from presenting them, we suggest alternatives from the MMM based on the personal face-to-face evaluations of individuals.

We provide mid-term alternatives to the MMM (IV).

- (I) The university's financing system, depending on the local government
- (II) The student evaluation exam to access a public university
- (III) The National Agency of Quality and Evaluation of Research Criteria
- (IV) Mid-term alternatives to the MMM

(I) The financing of a given university depends on the local government, and is closely linked with the number of students who graduate every year insofar as the more graduated students, the more financial resources for a university. Thus, according to Campbell's laws, the financing criterion is corrupted by the political ambitions of rectors, whose management is conducted by monitoring wrong university rankings. The direct consequence of this financing system is that learning processes lack rigor due to the progressive lowering of expected levels, and the removal of the effort-based system in HEIs. This fact deteriorates university students' knowledge levels every year.

An alternative to the current MMM would consist of changing the criterion for financing universities adopted by the local government by making it independent of the number of students who graduate. There is a plurality of options to avoid the metric "number of graduated students". Unfortunately, it depends on the local government, which also makes the mistakes of the MMM at the government level. A reasonable alternative financing criterion could be linked with the number of posts offered to students because it is clear that expenses are related to the number of students.

Independently of the metrics used by the government to finance HEIs, rectors have the possibility of not misleading the financing criterion "number of graduates" and of preserving the level of effort required of students, but this does not happen. Indeed, in Spain, rectors are politicians who are elected like professional politicians, and the only difference lies in the electoral body being made up of all university members, including students, instead of citizens.

The effect of the "number of graduated" metric is questionable in line with Campbell's laws, which essentially regard potential wrong human behavior. Rules and laws should take into account human weakness and attempt to avoid misleading human interpretations. It is understandable that rectors make efforts to increase their university's budget, which is probably not sufficient, but the financing criterion should avoid rectors feeling tempted to increase income by lowering the quality level expected of students.

Another alternative to the MMM is to delay questionnaires for teacher evaluations until at least 1 year after students were taught by lecturers for the last time. Generally speaking, the teaching performance bias relies on the effort required of lecturers to teach students. Hence, the lecturers who apply an effort-based system are penalized with the poor evaluations that their students score them with. There is a clear conflict of interest in students' opinions of teaching quality while the course is underway. It is not necessary to carry out an experiment, but only to know human passions [63].

(II) Another metric indicator to be reviewed in the Spanish HE system is the composition of the grade that allows students to enroll for a bachelor's degree in a public university. Presently, this degree is based on two components: the mark obtained by students in the last 2 years of high school and the mark obtained by the compulsory exam to gain access to a university.

The exam is a common one for all students in a given political district, independently of the bachelor's degree that students wish to study. It does not matter if this is civil engineering or foreign languages and literature. This fact, in theory, should influence the type of high school studies, but what actually occurs is the following: to be scored a better historic mark, as the exam to gain access to university is nonspecific, students choose high school studies that allow higher marks. When students select a degree that requires, for instance, mathematics, students frequently do not choose the appropriate high school studies, but easier ones like humanities instead of scientific ones. Once again, Campbell's laws appear given the use of a wrong metrics indicator.

What would a possible alternative be? There are several: each degree may have a specific serious exam, or one can join several related degrees and apply for the same serious specific exam for several closely related degrees. In these exams, historic scores should have no influence or, if one wishes, historic scores should be appropriate for the high school.

However, it is nonsense to suggest that literature and foreign students sit the same exam as civil engineering students. To avoid student complaints, the exam sat to gain university access is so easy that 95% of students pass it every year.

In addition, the exam to gain access to university is essentially the same from one year to the next. Instead of high school teachers teaching the official program, they train their students to pass the exam to gain access to university. Why? Because this ensures the high school a good position in the ranking of local high schools.

(III) The third alternative concerns lecturers' evaluations to obtain accreditation at several permanent university levels. This is related to them occupying the tenure position or promoting a full professor level.

For the last two decades, the accreditation system has been based on candidates' mechanic and impersonal evaluations, with a number of items concerned mainly with the appropriate number of them in ranked journals [64].

Once the candidate achieves accreditation, this is a sufficiency certificate for a category, and the university where the candidate works selects a committee, which is suggested by the candidate him/herself. Typically, there is only one candidate. So, his/her promotion evidently takes place.

Where do problems arise? As no serious study is available of the papers co-authored by the candidate, there is never any guarantee given by anyone during the process that the candidate is even the real author of papers with three authors or more. A level of repetition of the same idea hidden in several publications is also frequently found, but will not be recognized without serious research, at least of a personal type [65]. It does not suffice to appeal to a journal's quality because when one manages a paper, the editorial team does not know if ideas are repeated in another paper submitted elsewhere. So, there is no reliability, not even about the candidate's real authorship.

The present metric's automatic process avoids an old problem that we have experienced in Spain for more than 30 years, where the body of full professors is limited, and the influence of some families acting like a "mafia" by means of a national random draw of the candidates who have met the requirements to form part of an evaluation committee. This system today incurs no risk of subjectivity because the size of the potential committee members is uncontrollable.

All systems are imperfect. However, since 1975, Campbell's laws have emphasised the risks of using quantitative methods to evaluate persons. Negative effects do not finish with a lack of morality or the reliability of a given candidate's real merits. The issue becomes more complex because the local incentives of universities themselves are addressed to improve the quantity of research production, which causes collateral damage like: not taking the research risk to guarantee productivity; teaching young researchers and PhD students these overproduction habits; disregarding the teaching activity because it is not regarded as a relevant item for promotion.

What is the alternative? This is yet again a matter of degrees and culture. The main issue is to play down rankings, which is *per se* a business [14] of little credibility. Yet it depends on not only leaders' behavior, but also on both leaders' and followers' morality [66]. When scholars are advised about metric excess, and despite admitting wrongful behavior, they still reply that they are prisoners because their research students will be evaluated by the wrong MMM. This position is not morally right because, if it is true that they cannot change the present situation, the situation will alter if they change their attitude. Collateral damage, such as a reduced risk of research projects, guaranteeing productivity, and setting a bad example for young colleagues, is important enough to deserve a change of attitude.

In an attempt to be constructive, what data-use practices should people cultivate so we can appropriately employ data?

Before answering this question, we should think about what is important to evaluate and then, what is countable, or not, because probably not everything that counts is countable. How should we count? For example, vocation, discipline, generosity, the capacity to generate eagerness, patience, stamina, tolerance, integrity, credibility, self-esteem, awakening illusion, encouraging people, public speaking, the capacity to assume risks and face challenges. Evaluating these features is difficult, but not to the extent that they cannot be taken into account during the evaluation process.

All of these suitable characteristic traits are important for performing the public jobs of lecturer and researcher, and can be identified only by the personal interaction and personal development of the candidate to be evaluated by an evaluation committee [63,67].

Experience is not only important, but is also a suitable characteristic that can be evaluated and measured by carefully measuring its reliability; for instance, the real degree of participation in a multiauthor publication, or the same idea being repeated in several publications [38,49,50]. It is possible to measure teaching and research experiences using data. Excellence is the habit of doing a task very well. The search for quality is linked with slowness, while quantity is usually coupled with speed and a low risk. Thus, summarizing data-use practices is profitable to evaluate, but needs to be conducted by taking care of true reliable data. In any case, experience is a minimum requirement.

The stimulus of quantities is damaging. However, personal evaluations are essential to evaluate a candidate to ask about the above-mentioned suitable uncountable characteristics.

In order to avoid today's number-swamped and audit-crazed academic life, an alternative system should stimulate both the data of countable items and great challenges, even though they are unsuccessful. There is an overdose of irrelevant repetitive production which, apart from all the above-mentioned drawbacks, also leads academics to suffer stress and avoids the sufficient contemplative life needed to make great achievements. The present ideology of quantification and speed drives mediocracy. Quantitative items should be upper-bound and stimulate great challenges.

It is important to point out that the MMM is not something abstract that is applied to a country with a specific cultural factor, which means that many circumstances can be shared by another country, but at different intensities. Some people not being ready to totally accept the criticism of the MMM because they feel they prisoners of the system shows that it all depends on the decisions of political and academic powers to some extent.

The possibility of changing the MMM depends on political power when it draws up public university budgets, or because political power manages the National Agency of Evaluation of Research, even in some countries like Spain, where the university rector is also politically elected by the university community.

Political decisions do not depend only on rationale. The second level of influence stems from the academic power inside universities, and not by stimulating the quantitative issues of the MMM or lowering its degree of intensity. If the two first scenarios cannot be changed, i.e., financing comes from the state or a local government, and from the university's own policies, then one may resort to individual rebellion where each individual scholar can change his/her behavior with regard to their influence. Individual changes will resist through colleagues' laziness, interests, vanity or lack of courage, but the contagion effect produced by convinced individuals is not useless at all, mainly when important researchers support the strategy. Each individual scholar's individual initiative impacts both teaching and research because of the contagion effect. The hardest obstacle is starting the fight against the MMM and diffusing the social damage.

(IV) The following alternatives are addressed to discuss the potential changes in the paradigms of quantification by qualification: the amount of productivity items and their relevance; the speed achieved by slow behavior; pay-for-performance and intrinsic motivation; appreciating multidisciplinary activities without disregarding the quality of monodisciplinary activities. How can these ideas be implemented?

With a change in paradigm, the important issue is not how much or how many, but why and for what. A potential candidate to be promoted at a university should successfully answer in person these questions:

- (i) When and why did you decide to become a teacher?
- (ii) What important contribution, and why, do you make in the field?
- (iii) What are the purposes for you giving a class and how do you motivate students?

Questions (i) and (iii) are driven to identify the intangible suitable character capacities related to vocation, student motivation, courage, stamina, generosity, public communication, etc., that are countable by metric indicators. These issues are related to teaching and, of course, we should add the quantifiable items involved in the MMM, such as the years of teaching experience and degree of satisfaction, as well as relevant teaching publications. Issues related to (i) and (ii) cannot be undervalued because they are essential and have a strong impact on all students and the self-effort culture instead of on protection and coddling.

Question (ii) is extremely important for changing the mindset of quantity and speed to a suitable one for the relevance, immense challenge and interest of research. Appealing to the number of publications in a list of journals, which has been the case to date, is simply not enough.

The mere effect from candidates answering questions about these issues already has a clear impact. The expected potential claim made after any change, and after years of applying the MMM, is expected and deserves an answer. In fact, how objective is the evaluation of these unquantifiable issues?

The answer lies in the evaluation committee's qualified career and experience that deserve credit within human limits.

Finally, research experience needs to be considered, but also teaching activity, and only as a minimum sufficient threshold, by justifying both achievements and relevance. The minimum threshold may differ for each category, and is properly bound and weighted. The difference with the current MMM approach is that credit is not now deposited in some impact ranking of journals, but relevance must be proved and, of course, the repetition of the same idea in different papers must disappear, which is an essential issue recommended in [50]. This means that the quality of a paper is not automatically granted by a ranking of the journal where the paper is published.

Once the relevance of a publication is personally justified in the presence of a committee, clearly many current publications would disappear because this would be the way that most of an author's quoted misconduct would be made transparent. It is very different to show a list of publications where nobody pays attention to justify their relevance and innovation to a committee.

We believe that evaluations in universities need to change urgently and we hope that some ideas in this paper can help to improve them.

## 5. Conclusions

This paper analyzes the drawbacks of the MMM method and its widespread extension in the Spanish public education sector, starting with the university system. An explanation of the origin, rise and questioning of the MMM in the US army is included. The amateurism of political managers of the Spanish public system and their dependence on international rankings as a vector driver is leading the Spanish public system (both research and education) to continuous deterioration, which renders the Spanish Public system at risk of becoming unsustainable, as observed by the significant and continuous growth of the private university student numbers in Spain [23], while the public universities' student numbers oscillate according to the country's economy behavior.

In addition, the negative effects of the MMM from the mental health point of view appear due to the overpressure linked with rankings. In fact, 89.2% of teachers and researchers in Spanish public universities declared anxiety and burnout [46]. As university management adds other pressures due to increased bureaucracy [60] and student misconduct, the teacher and researcher jobs in the public university sector are becoming less attractive.

As we pointed out before, the effect of the MMM on the university lowers the level demanded of students. This also produces a contagion at high school by propagating the minimal effort culture in teaching centers in combination with overprotection, which makes young people become familiarized with victimization and fragility. As they are not trained to make the effort, they feel fragility with regard to the normal difficulties of real life: emotional shocks, labor problems, little self-esteem, low resilience, and being made victims of anxiety, social phobias and depression. In 2017, about 30% of Spanish young people suffered some mental illness, and 11% of young people aged 16–30 years had experienced a



panic crisis, stress, anxiety or social phobias [54]. The problem of fragility and weakness of character is not exclusive to Spain, although the extremely high unemployment rates for young people in Spain, above 25%, does not make this problem any easier.

A recent study shows that the young people's fragility problem also occurs in university students in the USA [68]. Cultural reasons favor the MMM being extended and diffused. One of the main targets of this paper is to warn about and prevent the risk of the MMM spreading to other environments, such as the Spanish public health system or its security system. The fact that rankings are commercialized [14] encourages managers to look at the mirror of rankings because it is an easy way to appeal to success, while in fact the quality of teaching and research worsens.

Although a country's culture is important, for instance, the way a university budget is drawn up and university managers are elected, or how students access a university and how it is regulated, it is quite clear that the majority of the MMM drawbacks are partially replicated in the public sector of most countries in the western orbit.

**Author Contributions:** This article was originally conceived and designed by L.J. and E.D.I.P. These two authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Spinoza, B. *Ética*; Espuela de Plata: Sevilla, Spain, 2017.
2. Damasio, A. *El Extraño Orden de las Cosas*; Planeta: Barcelona, Spain, 2008.
3. Veblen, T. *The Leisure Class*; Mentor Book: New York, NY, USA, 1912.
4. Drucker, P.F. *The Practice of Management*; Harper Collins: New York, NY, USA, 1986.
5. Liefner, I. Funding, resource allocation, and performance in higher education systems. *High. Educ.* **2003**, *46*, 469–489. [[CrossRef](#)]
6. Collini, S. *What are Universities for?* Verso: London, UK, 2017.
7. Froumin, I.; Lisyutkin, M. Excellence-driven policies and initiatives in the context of Bologna process: Rationale, design, implementation and outcome. In *The European Higher Education Area*; Curaj, A., Matei, L., Pricopie, R., Salmi, J., Scott, P., Eds.; Springer International Publishing: New York, NY, USA, 2015; pp. 249–265.
8. Boudard, E.; Westerheijden, D.F. France: Initiatives for Excellence. In *Policy Analysis of Structural Reforms in Higher Education*; Palgrave Studies in Global Higher Education; Palgrave Macmillan: Cham, Switzerland, 2017. [[CrossRef](#)]
9. Tsvetkova, E.; Lomer, S. Academic excellence as “competitiveness enhancement” in Russian higher education. *Int. J. Comp. Educ. Dev.* **2019**, *21*, 127–144. [[CrossRef](#)]
10. Hedges, C. Empire of illusion. In *The End of Literacy and the Triumph of Spectacle*; Amazon: Vintage, CA, Canada, 2010.
11. Denault, A. *Mediocracia: Como los Mediocres Llegan al Poder*; Turner: Madrid, Spain, 2019.
12. Delgado, L. Higher Education Systems and Institutions, Spain. In *Encyclopedia of International Higher Education Systems and Institutions*; Shin, J.C., Teixeira, P., Eds.; Springer Science + Business Media: Dordrecht, UK, 2018.
13. Jódar, L. Why the Spanish Public System is Wrong. In *Modeling Human Behavior: Individuals and Organizations*; Jódar, L., De la Poza, E., Merello, P., Acedo, L., Eds.; Nova Publishers: New York, NY, USA, 2016.
14. Becerra, J. El Negocio de los Rankings Internacionales: Así se Forran Gracias a la Guerra Entre Universidades. *El Mundo*, 2 October 2019.
15. Hazelkorn, E. *Global Rankings and the Geopolitics of Higher Education*; Imprint Routledge: London, UK, 2016. [[CrossRef](#)]
16. Pietrucha, J. Country-specific determinants of world university rankings. *Scientometrics* **2018**, *114*, 1129–1139. [[CrossRef](#)] [[PubMed](#)]
17. Olcay, G.A.; Bulu, M. Is measuring the knowledge creation of universities possible?: A review of university rankings. *Technol. Forecast. Soc. Chang.* **2017**, *123*, 153–160. [[CrossRef](#)]

18. Casani, F.; De Filippo, D.; García-Zorita, C.; Sanz-Casado, E. Public versus private universities: Assessment of research performance; case study of the Spanish university system. *Res. Eval.* **2014**, *23*, 48–61. [[CrossRef](#)]
19. De la Torre, E.M.; Gómez-Sancho, J.; Perez-Esparrells, C. Comparing university performance by legal status: A Malmquist-type index approach for the case of the Spanish higher education system. *Tert. Educ. Manag.* **2017**, *23*, 206–221. [[CrossRef](#)]
20. Revel, J.F. *El Conocimiento Inútil*; Planeta: Barcelona, Spain, 1989.
21. Teixeira, P.; Landoni, P.; Gilani, Z. *Rethinking the Public-private Mix in Higher Education*; Sense Publishers: Rotterdam, The Netherlands, 2017.
22. Levy, D.C. Global private higher education: An empirical profile of its size and geographical shape. *High. Educ.* **2018**, *76*, 701–715. [[CrossRef](#)]
23. Rodríguez, J.C. Las Cifras de la Universidad Pública y la Universidad Privada en España: 1983–2016. *Studia Siglo XXI*, Universidad (blog). 16 November 2017. Available online: <https://www.universidadsi.es/las-cifras-la-universidad-publica-la-universidad-privada-espana-1983-2016/> (accessed on 20 February 2020).
24. Merry, S.E. Measuring the World. Indicators, Human Rights and Human Governance. *Curr. Antropol.* **2011**, *52*, 83–95. [[CrossRef](#)]
25. Merry, S.E. *The Seductions of Quantifications*; The University of Chicago Press: Chicago, IL, USA, 2016.
26. Porter, T. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*; Princeton Univ. Press: Princeton, NJ, USA, 1996.
27. Muller, J.Z. *The Tyranny of Metrics*; Princeton Univ. Press: Princeton, NJ, USA, 2018.
28. Campbell, D.T. Assessing the impact of planned social change. *Eval. Program Plan.* **1979**, *2*, 67–90. [[CrossRef](#)]
29. Chrystal, K.A.; Mizen, P.C. *Goodhart's Law: Its Origins, Meaning and Implications for Monetary Policy*; Bank of England: London, UK, November 2001.
30. Fire, M.; Guestrin, C. Over-optimization of academic publishing metrics: Observing Goodhart's Law in action. *GigaScience* **2019**, *8*, giz053. [[CrossRef](#)]
31. Merton, R.K. La sociología de la ciencia: Investigaciones teóricas y empíricas. In *Recopilación e Introducción de Norman W. Storer*; Alianza: Madrid, Spain, 1973.
32. Lewis, M. More than Meets the Eye: The Under Side of the Corporate Culture of Higher Education and Possibilities for a New Feminist Critique. *J. Curric. Theor.* **2005**, *21*, 7–25.
33. Berg, B.; Seeber, B.K. *The Slow Professor*; The Univ. of Toronto Press: Toronto, ON, Canada, 2016.
34. Peters, T.; Austin, N. *A Passion for Excellence—The Leadership Difference*; Random House: Barnegat, NJ, USA, 1985.
35. Ravetz, J. *Scientific Knowledge and its Social Problems*; Transactions Publishers: New Brunswick, NJ, USA, 1971.
36. Biagioli, M. Watch out for cheats in citation game. *Nature* **2016**, *535*, 201. [[CrossRef](#)] [[PubMed](#)]
37. Manheim, D. Overpowered Metrics Eat Underspecific Goals. *RIBBONFARM*, 29 September 2016.
38. Fanelli, D. How many scientists fabricate and falsify research: A systematic review and meta-analysis of survey data. *PLoS ONE* **2009**, *4*, e5738. [[CrossRef](#)] [[PubMed](#)]
39. Morin, E. *La Mente Bien Ordenada*; Seix Barral: Barcelona, Spain, 2000.
40. Guinzburg, I. *El Talmud*; Berbera Editores S.A. de C.V.: México City, Mexico, 1980.
41. Raafat, R.M.; Charter, N.; Frith, C. Herding in humans. *Trends Cogn. Sci.* **2009**, *13*, 420–428. [[CrossRef](#)]
42. Christakis, N.A.; Fowler, J.M. *Connected*; Little, Brown and Company: New York, NY, USA, 2009.
43. Cubberley, E.P. *Public School Administration*; Houghton Mifflin Co.: Boston, MA, USA, 2010.
44. Luttwak, E.N. *Strategy: The Logic of War and Peace*; Harvard Univ. Press: Cambridge, MA, USA, 2002.
45. McNamara, R. In *Retrospect: The Tragedy and Lessons of Vietnam*; Vintage Books: New York, NY, USA, 1995.
46. Goyanes, M.; Rodríguez Gómez, E.F. Why do we publish? Prevalence, motivations and consequences of publish or perish. *El Profesional de la Información* **2018**, *27*, 548–558. [[CrossRef](#)]
47. Buela-Casal, G.; Guillen-Riquelme, A.; Diaz-Roman, A. Research Ranking of Spanish Public Universities. *Psicothema* **2019**, *31*, 351–362. [[CrossRef](#)]
48. Taleb, N.N. *The Black Swan*; Random House: New York, NY, USA, 2007.
49. Biswanger, M. Excellence by Nonsense: The Competition for Publication in Modern Science. *Open. Sci.* **2013**, *49–72*. [[CrossRef](#)]
50. Ramón y Cajal, S. *Reglas y Consejos Sobre la Investigación Científica*; Austral: Madrid, Spain, 2011.
51. Burke-Smalley, L.A.; Rau, L.R.; Neely, A.R.; Evans, W.R. Factors perpetuating the research-teaching gap in management: A review and propositions. *Int. J. Manag. Educ.* **2017**, *15*, 501–512. [[CrossRef](#)]

52. Van-Dalen, H.P.; Henkens, K. Intended and unintended consequences of a publish-or-perish- culture: A worldwide survey. *J. Assoc. Inf. Sci. Technol.* **2012**, *63*, 1282–1293. [[CrossRef](#)]
53. Aristóteles. *Ética a Nicómaco*; Alianza: Madrid, Spain, 2014.
54. Barómetro Juvenil de Vida y Salud 2017, Fundación Reina Sofia sobre Adolescencia y Juventud, Fundación de Ayuda a La Drogadicción. Available online: <http://www.proyectoscopio.es/barometro/barometro-juvenil-de-vida-y-salud-2017> (accessed on 20 February 2020).
55. Bell, A.S.; Rajendran, D.; Theiler, S. Job Stress, Wellbeing, Work-Life Balance and Work-Life Conflict Among Australian Academics. *Electron. J. Appl. Psychol.* **2012**, *8*, 25–37. [[CrossRef](#)]
56. Abramo, G.; Cicero, T.; D’Angelo, C.A. The dangers of performance-based research funding in non-competitive higher education systems. *Scientometrics* **2011**, *87*, 641–654. [[CrossRef](#)]
57. McIntosh, S.; McKinley, J.; Milligan, L.O.; Mikolajewska, A. Issues of (in) visibility and compromise in academic work in UK universities. *Stud. High. Educ.* **2019**. [[CrossRef](#)]
58. Dearden, J.; Grewal, R.; Lilien, G.L. Strategic Manipulation of University Rankings, the Prestige Effect and Student University Choice. *J. Mark. Res.* **2019**, *46*, 691–707. [[CrossRef](#)]
59. Carson, L.; Bartneck, C.; Voges, K. Over competitiveness in academia: A literature review. *Disruptive Sci. Technol.* **2013**, *1*, 183–190. [[CrossRef](#)]
60. *Ley 9/2017, 8 de Noviembre, de Contratos del Sector Público*; B.O.E. 272 de 9/11/2017; B.O.E.: Madrid, Spain, 2017. (In Spanish)
61. B.O.E. *de 3 de Enero de 2006, Resolución de 30 de Diciembre de 2005*; B.O.E.: Madrid, Spain, 2005; pp. 199–348. (In Spanish)
62. Luttwak, E.N. *The Pentagon and the Art of War*; Simon and Schuster: New York, NY, USA, 1986.
63. Kallio, K.; Kallio, T.J.; Tienari, J.; Hyvönen, T. Ethos at stake: Performance management and academic work in universities. *Hum. Relat.* **2016**, *69*, 685–709. [[CrossRef](#)]
64. Albert, C.; Davia, M.A.; Legazpe, N. Determinants of Research Productivity in Spanish Academia. *Eur. J. Educ.* **2016**, *51*, 535–549. [[CrossRef](#)]
65. Lee, I. Publish or perish: The myth and reality of academic publishing. *Lang. Teach.* **2014**, *47*, 250–261. [[CrossRef](#)]
66. Bandura, A. *Moral Disengagement*; Worth Publishers: New York, NY, USA, 2016.
67. Wessels, J.S. The human spirit as a necessary condition for higher education: A risk assessment. *J. Transdiscipl. Res. S. Afr.* **2015**, *11*, 13–28. [[CrossRef](#)]
68. Haidt, J.; Lukianoff, G. *The Coddling of the American Mind*; Penguin Press: New York, NY, USA, 2018.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).