Where is the theory of business school program rankings? A critical assessment.

Daniel de Magalhães Chada[1,2], Alexandre Linhares[2]

1. Department of Psychology, University of California, Berkeley, danielchada@berkeley.edu.

2. The Brazilian School of Business and Public Administration, Getulio Vargas Foundation, P. Botafogo 190, Rio de Janeiro 22250-000, Brazil, alexandre.linhares@fgv.br

Abstract

Business school rankings exert an oversized influence on schools' strategies, programs, and operations. In this paper we take a step back and question the philosophy and the methodology involved in such rankings, and we argue that rankings fall prey into two serious problems: i) rank anomalies, and ii) bifurcations. Our first study considers clustering of the Financial Times dataset, and comparing the clusters to the presented ranking–which shows a large number of anomalies. Our second study introduces a new statistical method and demonstrates that the Business Week ranking of top US-based MBA programs shows the presence of bifurcations. The linear structures imposed by rankings cannot capture the multidimensionality of the space in which Business Schools compete. We hope scholars will place increased scrutiny in the philosophical question: where is the theory and the scientific validity of rankings?

Keywords: Business School Rankings; Computational Cognitive Science; Bayesian Reasoning; Exploratory Statistical Analysis; Massively Multidimensional Spaces

"*We're going to start rating colleges not just by which college is the most selective, not just by which college is the most expensive, not just by which college has the nicest facilities–you can get all of that on the existing rating systems. What we want to do is rate them on who's offering the best value so students and taxpayers get a bigger bang for their buck,*"

—President Barack Obama, Speech at the University of Buffalo, 2013.

1. Introduction

Business schools programs are regularly ranked by Business Week, US News & World Report, The Economist, Fortune, Financial Times, the Wall Street Journal, amongst other organizations and periodicals. Their publication, beginning in the 1980s, have generated high controversy in the U.S. These rankings exert deep influence across the business school landscape (Pfeffer and Fong, 2004; Gioia et al., 2000; Gioia and Corley, 2002; Corley and Gioia, 2000; Alvesson, 1990). They directly affect the perceptions of current students, alumni and prospective students in regards to the quality of the ranked schools. Their influence has repercussions to the extent that schools alter their curricula, fire faculty, and adapt teaching methods with the explicit objective of rising in the ranks. Zell (2001) elaborates on this change of behavior since the rise of the business school rankings. Pfeffer and Fong (2004) explain how i) business schools taylor their curricula in attempts to rise on the ranking; ii) professors "dumb down" their courses in order to receive better reviews from students; iii) the press (rather than academia) has led the way in defining standards of world-class business education and iv) the above points cause a standardization of business schools, which is detrimental to students (who lose options for different types of education) and to schools. Corley and Gioia (2000) explain how "the rankings by these magazines have come to dominate many business schools' sense-making and action-taking efforts". Business Week, specifically, calculates a Return on Investment (ROI), in order to measure to what extent alumni have achieved financial success, and how quickly they have done so—a narrow focus which can be arguably detrimental to the long-term perspective.

Dichev (1999) questions the validity of rankings as a whole, concluding from a cross-rankings correlation that neither the Business Week nor the U.S. News rankings "should be interpreted as a broad measure of school quality and performance", and that the "absence of positive correlation combined with reversibility in changes implies that one should avoid a broad interpretation of the rankings as measures of the unobservable 'school quality"'. Still others suggest alternate evaluation methods for schools, using different indicators to provide a 'better' ranking system (Tracy and Waldfogel, 1997) or better principles (Cornelissen and Thorpe, 2002) in order to better reflect the qualities of each institution. (These, nevertheless, also impose the order structure, which is the critical point of focus here.)

While students and alumni generally regard the rankings as a valid metric of the quality and reputation of the schools, faculty and staff generally share a more adverse view of the ranking system. Among the latter they are viewed as poor quality indicators of the education provided by an institution, studies show that there is virtually no correlation between a position in the ranking and academic production Siemens et al. (2005); Trieschmann et al. (2000). Other evidence shows that both the rankings themselves Elsbach and Kramer (1996) and the changes caused by them Pfeffer and Fong (2004); Zell (2001) elicit responses ranging from mild annoyance to "outright rebelliousness" amongst faculty. As a testament to the power and influence of rankings, there is empirical evidence showing the correlation between the rankings and the resignations of the deans of schools that score poorly on them Fee et al. (2005). From this body of literature we may conclude that rankings hold a huge sway over institutions and their strategies—and over the students' choices concerning which one to attend—despite their clear dissociation from any objective measure of the relative quality disparity of education at each institution. Similar discussions arise from university–level rankings (Ehrenberg et al., 2001; Liu and Cheng, 2005; Florian, 2007; Ioannidis et al., 2007). This leads us to our first proposition:

Proposition 1. *Rankings directly and overwhelmingly influence a school's strategy and operation.*

Because the public at large–including the American President–seem to be taking rankings at face value, we feel that scrutiny and skepticism about the philosophy and methodology involved in program rankings is warranted.

Though we will focus on the Financial Times and Business Week rankings in this study, we will see that questions raised concern the underlying methodology of all school rankings–the first of which is the loss of crucial information. Rankings, by their mathematical nature, create serious anomalies—either by placing dissimilar schools in close rank positions, or by placing similar schools in far rank positions. The simplest such case happens with 4 schools and two (binary) dimensions, as we will see below.

1.1. The problem of rank anomalies

A ranking can be defined as the mathematical structure known as an *order*: given two distinct entities ϵ_1 and ϵ_2, the statement $\epsilon_1 \prec \epsilon_2$ denotes that ϵ_1 *precedes* ϵ_2, or ϵ_1 *dominates* ϵ_2. The stated meaning in a school ranking is that if school ϵ_1 precedes school ϵ_2, then, generally, ϵ_1 should be preferred to ϵ_2 by prospective students, by faculty in search of job positions, by potential employers of alumni, and by other observers and stakeholders—and the strong phrase, *"best schools"*, is explicitly used. An order, brought by a ranking, projects schools into a unidimensional, mathematically transitive space, in which there can be no ambiguity, circularity, or niches. Is this unidimensional, transitive, space the best domain to project business schools?

Consider a *Gedankenexperiment*. Imagine four business schools. Two business schools, *LE* and *LA*, are strictly concerned with theoretical development and blue-sky research. There is no concern with practicalities, hardly any focus at teaching, and case studies and examples are explicitly prohibited. There is one striking difference between the two: *LE* is an expensive school, while *LA* is an affordable school. There is absolutely no other difference between the schools: all professors, installations and every other imaginable characteristic are exactly the same. The other two business schools, *RA* and *RE* sharply focus on example after example, and never attempt to find generalities, similarities, analogies, or models that join characteristics or general ideas from even two individual examples from their vast libraries. The only difference between the two schools is that *RE* is expensive and *RA* is affordable.

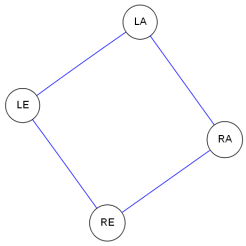


Figure 1: *Each of the four schools is related to two others by one—and only one—of their dimensions.*

What is the structure that captures the relationship between these schools? There are at least two equally plausible structures: a *grid*, or a *ring*:

- A *grid* structure has two axis x, y in which entities differ—rather like price versus quality, or height versus weight. In this case, the dimensions are (obviously) affordability-exclusivity and a fundamentalist focus on examples-theoretical constructs.

- The *ring* structure also suggests itself: note that, if one starts at any school and moves in either the clockwise or counterclockwise direction, one will rapidly find oneself at the beginning of the journey.

The *naturally intuitive* structure for the schools is either a grid or a ring. One can, of course, *project* these schools into an order, creating a ranking $\{\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4\}$. But that rank will not be its natural structure, as necessarily there will be schools that are placed next to one another while differing in all dimensions. The rank can respect school similarity (following the ring), or it can prioritize one dimension over another:

- If the rank follows the ring (clockwise or counterclockwise), the first school ϵ_1 will share *one* crucial dimension with ϵ_4 (just as it will with ϵ_2)—but ϵ_1 will share no dimensions with ϵ_3, the third-ranked 'opposite' school. Most importantly, this happens regardless of how the rank is construed. In other words, the first school will be significantly more similar to the last school than to the penultimate school (which, inconsistently, will be one position closer to the first in the ranking).

- If the rank prioritizes one dimension over another (i.e., following the grid structure), schools ϵ_2 and ϵ_3 will not share dimensions but will be next to each other in the rank. Students that strongly prefer school ϵ_2 but are accepted only by ϵ_3 and ϵ_4 face a hard prospect, as ϵ_3 will not share any dimension with their preference, while ϵ_4 will share one such dimension. Should students go to ϵ_4 to satisfy one of their preferences? In this case, they would risk the prejudice of the lowest ranked school in the whole set. Note that this occurs no matter which dimension is prioritized in the ranking's construction.

Proposition 2. *Projection into a unidimensional domain loses crucial information, as similarities between schools necessarily vanish.*

Schools can be close to each other in the ranking, but far from each other in their 'true nature'. On the other hand, schools can be far from each other in the ranking, and close to each other in their profile. Herein we define this as a *rank anomaly*. This article hopes to achieve two objectives. First, we would like to introduce a research method that can be widely applied to analyze social, organizational, and economic data. The second objective is to show the power of this method through the analysis of the data of Business Week's MBA program rankings. The results obtained demonstrate rank anomalies in the published rankings—providing an additional perspective for the critical literature of such rankings.

2. Rank anomalies: the Financial Times case

One would be inclined to believe that, say, school 7 and school 8, share similar data: that schools close in a ranking actually have a high probability of belonging to the same cluster. However, this is not the case.

These remarks do apply to all the Business Schools rankings we have tested. For reasons of space we have focused on Business Week case, but, for example, the Financial Times 2013 ranking of MBA programs own statistics clusters schools in a very poor relation to their ranking.

The Financial Times, though, notes that:

> Although the headline ranking figures show the changes in the survey year to year, the pattern of clustering among the schools is also significant. A total of 204 points separate the top school from the school at number 100 in the ranking. The top 10 schools, from Harvard Business School to University of Chicago: Booth, form the leading group of world-class business schools, separated by 82 points. The second group is topped by IE Business School, which scored 65 points more than Imperial College Business School, leader of the third group. The fourth group, which includes schools ranked from 74th to 100th, is headed by University of Cape Town GSB.

To which we respond that the FT is disrespecting sound mathematics. These numbers have dimensions associated with them, and one cannot simply sums number of different dimensions and proclaim that forms a cluster. This is analogous to dividing 1,000,000 Euros by 700 kilograms, or multiplying the result by 300 calories. This is not even undergraduate school mathematics, yet, by proclaiming that one can simply sum a "weighted salary" rank with an "alumni recommend" rank, each of which coming from a different statistical distribution and different mathematical dimension, the Financial Times seems to be fully exposing itself to mathematical folly. This leads us to a third proposition:

Proposition 3. *There is a large number of rank anomalies in the Financial Times Rankings. By "anomalies" we mean that schools that are very close in the ranking, if clustered by the underlying data, become increasingly far apart.*

Before engaging in an alternative visualization for rankings, we address the more general problem: *the imposition of structure*. Further on we address its impact in decision-making, exemplified by the choice of business school by an MBA candidate.

2.1. The imposition of structure

Structures are imposed by most analytical methods. Clustering methods will always find disjoint sets in data. Ranking (or order-based) methods project entities into a domain that must be isomorphic to either $\mathbb{N}$, $\mathbb{Z}$, or $\mathbb{R}$. Decision tree methods will create branchpoints to classify the data, and so forth.

Nature, on the other hand, is indifferent to our methods. Nature presents us with a bewildering array of different forms and structures—as do societies, firms, and other complex systems. Humans find structures by studying data and carefully comparing and contrasting this information to previously experienced structures Linhares and Freitas (2010); Linhares and Chada (2013). Our analytical methods, however, *impose* structures to data. This imposition can be harmful in a number of ways:

- It may suggest hypotheses which are not warranted. A ranking of living beings, *scala naturae* (or "the great chain of being"), was the unquestioned christian doctrine until Carl Linnaeus proposed the tree alternative; this "great chain of being" hypothesis—which goes upwards from rocks, plants, animals, man, spirit, angels and god—suggested a hierarchy of beings that proceeds towards 'greatness'; while the *tree of life hypothesis* suggests a common ancestor, speciation, and the exploitation of niches.

- Moreover, the imposition of structures may blind us to important relations hidden in the data. Prisoners

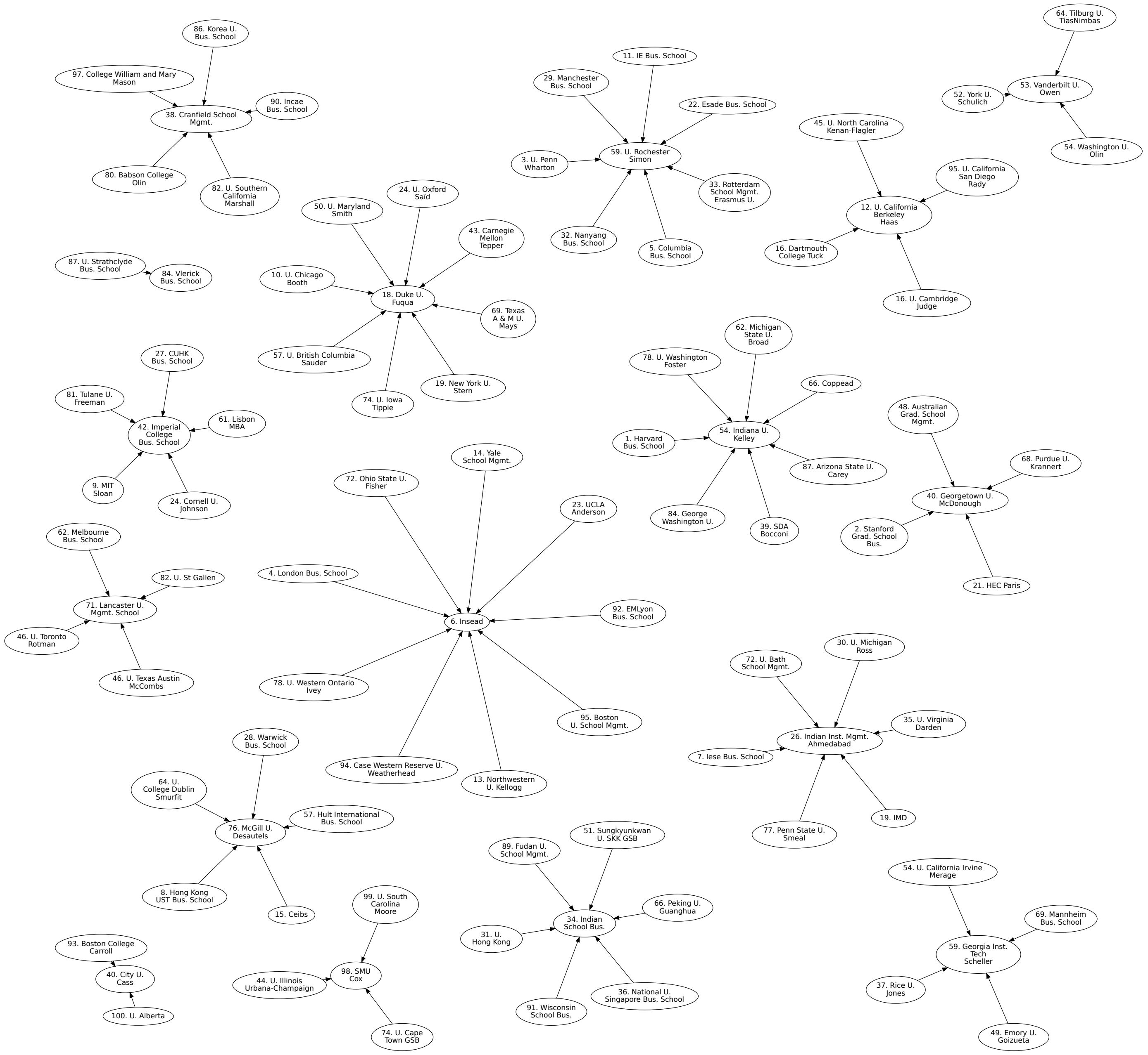


Figure 2: A clustering of the FT school programs dataset. In order to clusters the schools, we have used the Financial Times Ranking's data: *i) Salary today, ii) Weighted salary, iii) Salary percentage increase, iv) Value for money rank, v) Career progress rank, vi) Aims achieved rank, vii) Placement success rank, viii) Employed three months (%), ix) Employed three months (% reported by school), and x) Alumni recommend rank.* We have employed the *Jaccard distance* and the *affinity propagation* algorithm (Frey and Dueck, 2007)–one of the most successful modern clustering methods.

Figure 3: The Financial Times 2013 Ranking, superimposed over the clustering of their underlying data. This is the same graph of Fig. 2., with the addition of edges in red, representing the order of schools given by the Financial Times. One would expect that the data of the Financial Times Ranking would cluster schools in reflection to their position in the rankings. However, this is not the case–and there are serious rank anomalies found. Note that the rank anomalies are preserved in numerous statistical distance measures. Even for the authors of the present study, which have been focused on the philosophical and methodological problems of rankings for years, this visualization strikes us as astonishing. Though we had from the start expected that the tight clusterings obtained from ranking data would have had a loose coupling with the rank order, we do admit that we had never expected actual results to be as convoluted as in here.

generally self-organize into groups (gangs). Clustering is able to capture the increased intra-group interaction that dimensionality-reducing methods (such as χ^2 or the use of z-values) cannot. Ranking prisoners in order of 'violence propensity', or guards in terms of 'abuse of power propensity', will create the previously mentioned rank anomalies and will most likely neither reflect nor predict violence between individuals in any meaningful way. One needs to know how individuals interact, not how they rank in a single dimension.

- Finally, the structures may simply be inconsistent with the data, as in the case of rank anomalies. Do these anomalies appear in publicized rankings? If so, can we detect them? (As we will see below, in the "Top-30" Business Schools of America—according to Business Week—the answer to both questions is a resounding *yes*.)

There is, however, no need to presuppose a form when analyzing data. Cognitive scientists Charles Kemp and Joshua Tenenbaum Kemp and Tenenbaum (2008) have developed and algorithm enabling the *automatic discovery of form*. While Kemp and Tenenbaum have published their work *as a cognitive theory*, in this study we present their approach *as a new analytical method*, and we apply it to school rankings. We hereafter refer to the model we will use as the KT-Algorithm and its computed structures as KT-Structures–not to be confused with Hermitian structures(Basel, 2004). In the next section we summarize the main concepts of Kemp and Tenenbaum's mathematical model.

3. KT-Structures

Kemp and Tenenbaum (2008) have developed a model which, through hierarchical Bayesian inference, can explore and discover the underlying form that best adapts to a given dataset:

> Discovering the underlying structure of a set of entities is a fundamental challenge for scientists and children alike. Scientists may attempt to understand relationships between biological species or chemical elements, and children may attempt to understand relationships between category labels or the individuals in their social landscape, but both must solve problems at two distinct levels. The higher-level problem is to discover the form of the underlying structure. [...] the lower-level problem is to identify the instance of this form that best explains the available data. (p. 10687)

The authors provided, as a psychological theory, a method for the unsupervised learning of structure and form. We posit herein that this method can, moreover, be used towards data analysis. Statistical methods currently focus on the application and optimization of a given structure to data, while presupposing a specific underlying form, such as groupings (e.g., clustering), trees (e.g., hierarchical clustering, minimum spanning tree), or spacial representations (e.g., multidimensional scaling, self-organizing maps). That is, if one applies a clustering method to a set of data, one is assuming that clusters provide a suitable form to analyze and understand the data. If one applies decision trees, one is projecting that the data can be best understood as having no cycles. Similarly, a school ranking projects schools into a unidimensional, mathematically transitive, lens. The question we pose, therefore, is whether that is the best form to analyze the data provided and lead prospective students to optimal decisions concerning school choice.

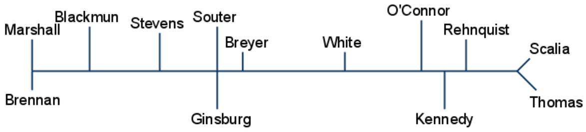
There are two important ideas involved in their model: i) the use of a hierarchical Bayesian method to analyze data; and ii) the use of graph grammars and graph redescriptions. Through simple operations based on graph-grammars and Bayesian inference, the algorithm is capable of exploring the space of possible forms and their instances to find a best-fit representation of the provided dataset. Let us look at each of these ideas in the following subsections.

Figure 4 reproduces two examples from the dataset used in (Kemp and Tenenbaum, 2008): The discovery of a chain that arranges U.S. supreme court justices (in 2008) from liberal (Marshall & Brennan) to conservative (Thomas & Scalia) in fig. 4a. This structure is extracted from a simple record of each justice's votes over approximately 1,500 cases. It is striking that the system is able to find that judges self-organize in a one-dimensional manner (reflecting 'conservative' and 'liberal' positions). In fig. 4b, from the number of results of Google^{TM} searches of the phrase "Xi told Xj", where X_i and X_j are different members of the Bush Cabinet, the system is able to find the correct form (a hierarchy), which is, again, the correct structure within this form.

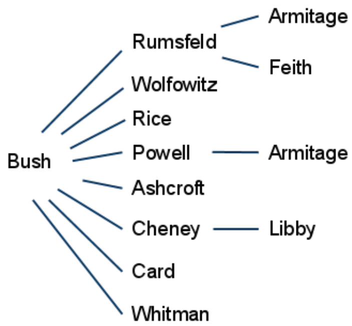
These examples illustrate the applicability of the algorithm towards exploratory data analysis: without knowledge of meaning or an *a priori* defined form, the algorithm arrives at a visualization that is not only intuitive but informative for those wishing to interpret the data.

3.1. Hierarchical Bayesian model: forms, structures, and data

Bayesian models have been applied with promising results throughout the literature Assaf and Josiassen (e.g., 2012); Assaf et al. (e.g., ming); Park and Kim (e.g., 2013); Smith et al. (e.g., 2000), and so have graph-theoretical models (e.g., M. Chirgui and Penard, 2011; Van Der Merwe et al., 2004). Examples of *hierarchical Bayesian models* can also be found stemming from areas as diverse as marketing (Abe, 2009) to political science (Lock and Gelman,



(a)



(b)

Figure 4: *(a) Spectrum extracted from justices' votes; (b) Hierarchy of the Bush cabinet.*

2010). But applied studies in which *graph structure is derived from the underlying data* remain lacking.

In Kemp and Tenenbaum's model, starting from dataset D, the algorithm attempts to find a form F and the structure S that best captures the relationships within the dataset. Input data may be expressed either as features and elements or as triangular relational matrices, containing data about the relations between items to be explored. This cognitive aspect of discovery occurs on different levels of abstraction concurrently. The possibilities are generated via graph-grammar splits (which we will see below) and the system seeks then to maximize the posterior probability:

$$P(S, F|D) \propto P(D|S)P(S|F)P(F)$$

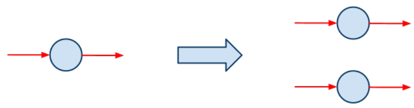
That is, what is the probability of a form F and a structure S, given a dataset D? In the hierarchical model, this probability is proportional to the product of i) the probability of dataset D given structure S, ii) the probability of a structure S given the form F, and iii) the probability of form F.

The algorithm has eight main forms that it explores: partition, chain, tree, hierarchy, ring, cluster, grid and cylinder.Initially $P(F)$ is given by a uniform distribution over all possible forms of the model.

$P(S|F)$ is given by the number of structures compatible with a given form:

$$P(S|F) \propto \begin{cases} \theta^S & \text{if } S \text{is compatible with } F \\ 0 & \text{otherwise,} \end{cases}$$

(a)



(b)

Figure 5: (a) Tree-generative Graph Operation. (b) Cluster-generative Graph Operation

i.e., if S is incompatible with F, then $P(S|F) = 0$. Otherwise it can be computed utilizing the Stirling number of the second kind and the number of k-cluster structures for a given form as normalizing constants, as described in (Kemp and Tenenbaum, 2008). Graphs with numerous clusters are penalized through parameter θ.

.The second important concept is the use of *graph grammars* to generate the forms and structures reflecting the data.

3.2. Graphs and graph grammars

Graph theory provides a mathematical framework to understand objects and their relations. The KT-Algorithm of (Kemp and Tenenbaum, 2008) defines the hypothesis space through graph operations and, through Bayesian inference, make use of these simple operations to generate a given structure as a possible fit to the data presented. These generating methods are graph grammars. Any particular structure S, that is, a specific instance of a form F (tree, ring, partition, etc.) can be generated by repeatedly applying a graph grammar operation over the graph of dataset D. To illustrate, consider the graph grammars for trees and chains:

i) Trees: Suppose all objects are put in a single cluster, C_1. A graph grammar for trees will select a subset of these objects to move to a new cluster C_2, and create a branch point $B_{\{C_1,C_2\}}$ that leads to C_1 and C_2, see Fig. 5a.

ii) Chains: Suppose, once again, that all objects are put in a single initial cluster C_1. A graph grammar for chains will create a cluster C_2, and split C_2 from C_1. No branchpoint is created in this form, of course, see Fig. 5b.

If the same operation is used on subsequent clusters, i.e., not only on a starting cluster with all objects contained therein. This enables the 'splitting' process to continue until a final structure is reached.

In the KT-algorithm a cluster is chosen to be split and two new clusters will replace it. Each element of the old cluster must be assigned to one of the new clusters. Two elements are chosen at random and placed in each of the new clusters, then the algorithm iterates each item randomly and makes a greedy assignment for each. Also, items may eventually be moved between clusters, if the model finds that this would create a structure that is more adequate to the data. This non-deterministic process dictates that the algorithm as a whole is non-deterministic.

We will close our description over the principal points of the original authors' algorithm by focusing on the process of generating the size of the split space via Stirling Numbers.

3.3. Stirling Numbers

As we have mentioned, the KT algorithm chooses, at each iteration, one best split operation to apply over the graph. This requires generating the possibility space of all potential split locations. In order to generate the number of possible splits, the original authors use the Stirling numbers of the second kind as a normalizing constant. A Stirling number is calculated as follows:

$$S(m,n) = \frac{1}{n!}\sum(-1)^k \binom{n}{n-k}(n-k)^m$$

This shows the number of ways of distributing m distinct objects into n identical containers with no container left empty. We refer the interested reader to Kemp and Tenenbaum for further details.

Given this brief summary of the KT-Algorithm, we may now proceed to apply the model to the US Rankings of US-based MBA Programs.

4. Experiments with KT-Structures

4.1. The BusinessWeek Ranking

Our experiment computes KT-Structures of the BusinessWeek 2008 ranking. The purpose is to compare and contrast the rankings widely used with the KT-Algorithm. We computed all the 21 possible forms provided in the method, and we concentrated attention to those that suggested rank anomalies. Of these, the tree and hierarchy structures provided high posterior values and readily presented potential rank anomalies, and we concentrate focus on them here.

4.1.1. Materials and Methods

We use the data provided in the Business Week MBA program ranking, and we computed all possible KT-Structures. We ignored the values in the fields "2006 Rank" and "2008 Rank", as we do not want to skew the results towards those generated by Business Week—had we included such dimensions, the correlation between ranking distance and KT-structure distance would become artificially inflated. In the remaining dataset, there are 12 variables, and nothing beyond those values is assumed to either exist or have any importance. These dimensions are: graduate poll, corporate poll, intellectual capital, tuition and fees, pre-MBA pay, post-MBA pay, selectivity, job offers, general management, analysis, teaching, and careers. These last four dimensions ranged from A+ to C, and we changed these results to numerical values, that is, A+, A, B, and C were translated to 1, 2, 3, and 4, respectively. This maintains the sequential nature of the scores – which would be lost with dummy coding, *per se*.

There is an important aspect to be noticed here. The model does not know that an A-grade is better than a C-grade (if anything, 4 is higher than 1), or that a higher post-MBA pay value is better than a lower one. The method does not have any information concerning the meaning of all these variables. But there are strong relations between the data: ranks are provided by orders whereas tuition, fees and pay are determined by the market, grades are obtained through Business Week's polls, and so on. The model is able to compute the structures based only on the underlying data, and does not need to understand the *meaning* imbued in each dimension—the problem of meaning is far from solved in cognitive science (Linhares, 2000; Linhares and Brum, 2007; Hofstadter and FARG, 1995).

4.1.2. Numerical results

The most interesting computed form is a hierarchy; presented in Figure 7. At a macro level, this form has some semblance with the original ranking. We define distance between two schools (i, j) is measured by counting the number of edges from the origin school's cluster to the destination school. The rank distance, on the other hand, is simply obtained by $|r_i - r_j|$, where $r_n$is the position of school n in the rank. Note that the domains are quite distinct, as distances in the KT-Structure tend to be smaller, yet, there is positive correlation between the rank and the hierarchy ($r = .65$—and covariance is 8.34), as can be seen in Figure 6. This shows that—at a large scale—there is some agreement between the rank and the KT-Structure.

The striking characteristic of trees and hierarchies—as contrasted to rankings—is the possibility of branch-points. If the reader will allow a metaphor: given the data, schools are better viewed as cities organized alongside a river rather than as elevator stops on a skyscraper. The glacier melts at the left side of figure 7, with the cluster comprising Harvard, Stanford and Wharton. As one moves downstream, the differentiating variable (at this point) is post-MBA pay: the first cluster with three schools are the only ones over \$120k, the second cluster with values ranging from \$105k (Chicago) to \$116k (MIT). The cluster comprising Michigan (\$105k) and Duke (\$100k) is followed by one comprising Cornell (\$96k) and NYU (\$95k). There are three schools downstream with post-MBA pay of \$100k or more (UCLA, Virginia, and CMU), but at this stage many other variables become increasingly relevant, and the tree branches.

A small stream leads to Yale, Maryland, and Olin. A combination of relatively undesirable data explains this cluster: the schools share C's in "general management" and "analysis" (and B's in "careers"), they are low-ranked in the corporate poll (positions 33, 41, and 42), and they

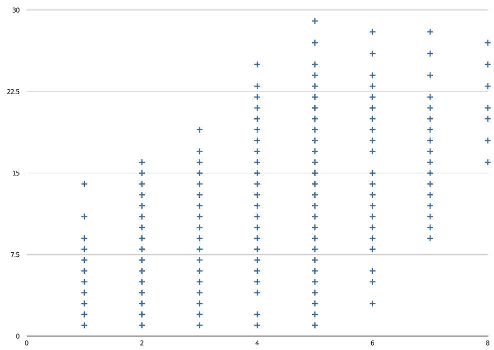


Figure 6: *The KT-Structure distance between schools plotted against their ranked distance.*

are relatively expensive. These traits lead us to interesting distortions between this tree and the rankings.

These are serious rank anomalies, and we turn our attention to the underlying data analysis next.

4.1.3. Analyzing school distance metrics

In this section, we focus on schools ranked #22 (Brigham Young) to #29 (Georgia Tech), in order to better understand the KT-algorithm's results. Let us observe how the schools in this set are different from each other, through the lens of different distance metrics. Let $M_{\{22,\dots 29\}}$ be the matrix with the 8 schools in question, and the 12 dimensions used by BusinessWeek. We can now measure school similarity through different 8×8 distance matrices:

- Let M_A be the resulting matrix, given by the *Bray-Curtis distance:* $d(u, v) = \frac{\sum_i u_i - v_i}{\sum_i u_i + v_i}$.
- Let M_B be the resulting matrix, given by the *squared Euclidean distance:* $d(u, v) = ||u - v||_2^2$.
- Let M_C be the resulting matrix, given by the *city block distance:* $|x1 - x2| + |y1 - y2|$.
- Let M_D be the resulting matrix, given by the *Chebychev distance:* $d(u, v) = max_i |u_i - v_i|$.

For plotting purposes, we normalize these distance matrices to values over the set [1, 2]; the resulting distance graphs can be seen in Figure 9.

It is particularly striking to note the maximum distance between #22 (Brigham Young) and #24 (Yale) under each of these distance metrics. Given that these schools stand merely 2 rank positions apart from each other, one would expect that, given the data used by BusinessWeek, they would not be far from each other, and certainly not the farthest schools among the group. Yet, that is precisely the case: under Chebychev, under squared Euclidean, under City Block and under Bray-Curtis distance metrics, Brigham Young and Yale are consistently *the farthest schools apart.* Given that this effect might be exclusive to the 4 distance metrics used above, we have also measured the Canberra distance, the correlation distance, the cosine distance, the Euclidean distance, the Hamming distance, the Jaccard distance, and the Minkowski distance (with $p = 1$) metrics. The same effect is found in all of them.Quite surprisingly, schools #22 and #24, as ranked by BusinessWeek, do not share many similarities.Another crucial point here is that the schools in the red and blue subsets show a tendency in all visualizations to be closer to same-color counterparts than from those of the other color, despite the fact that this does not respect the ranking order. This clustering reflects and corroborates the bifurcations discovered by the KT-Algorithm.

4.1.4. Discussion: a student's perspective

As a demonstration of the explaining power of the KT-algorithm, consider the following example. Suppose a student preferred the University of Washington's Foster School (no. 27), but was rejected there and accepted by two schools: Yale (no. 24) and Georgia Tech (no. 29). The student's choice seems easy, as Yale is no doubt better ranked.

The KT-Structure, however, tells a different story, placing Yale far from the student's preferred Foster. Here is why. If the student chooses Georgia Tech, tuition costs drop slightly from Foster's \$64,902 to Georgia Tech's \$64,152–while Yale will charge \$93,098. If the student choses Georgia Tech, Foster's "B" in "general management" is also found in Georgia Tech–while Yale holds a "C". If the student chooses Georgia Tech, Foster's "B" in "analysis" is reflected by an "A" in Georgia Tech's grade–while Yale holds a "C". Georgia Tech, at the 28th position in the corporate poll, is much closer to the preferred Foster's 26th position than Yale (33th position).

Of course, by choosing Yale over Georgia Tech, there are also significant gains in other dimensions. Yet *these gains are in dimensions that the student did not prioritize by expressing a preference towards Foster.* The ranking keeps moving further away from the student's preferred school characteristics. The preferred school held the 30th position in the graduate poll; Georgia Tech holds the 31st—but Yale is at the 19th position. In "intellectual capital", the preferred school held the 29th position, Georgia Tech holds the 26th position—but Yale is number 10. In school selectivity (perhaps a minor concern to our already accepted student), the preferred school accepts 30% of applicants, Georgia Tech accepts 29%—while Yale is much more selective, at 14%.

Of the 12 dimensions considered in building the ranking, Yale differs significantly in 7 dimensions from both the student's preferred school and from Georgia Tech (and also from Brigham Young). This is why the KT-Structure places schools like Maryland (26) close to Washington University's Olin (28), while both are far from the University of Washington's Foster (27) and Georgia Tech (29) (which also resemble each other in many dimensions). Instead of differentiating them, the rank alternates between these

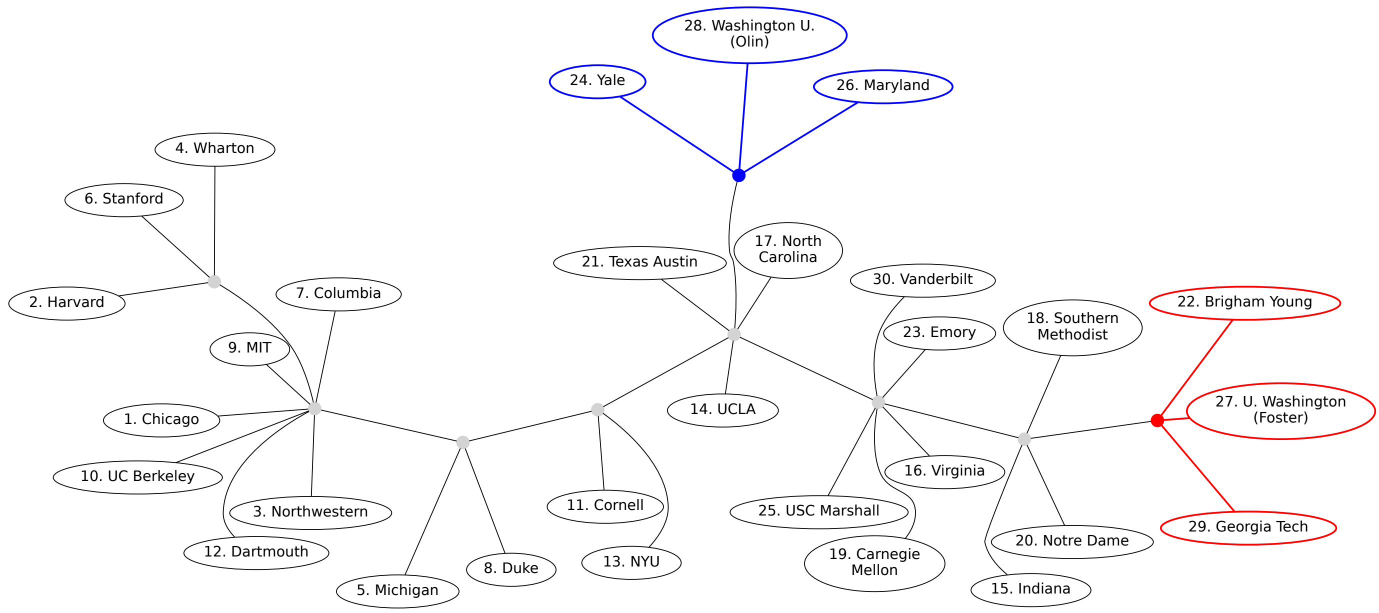


Figure 7: *The generated KT-Structure: an undirected hierarchy with no self-links. We emphasize two sets of schools in red and in blue for further study.*

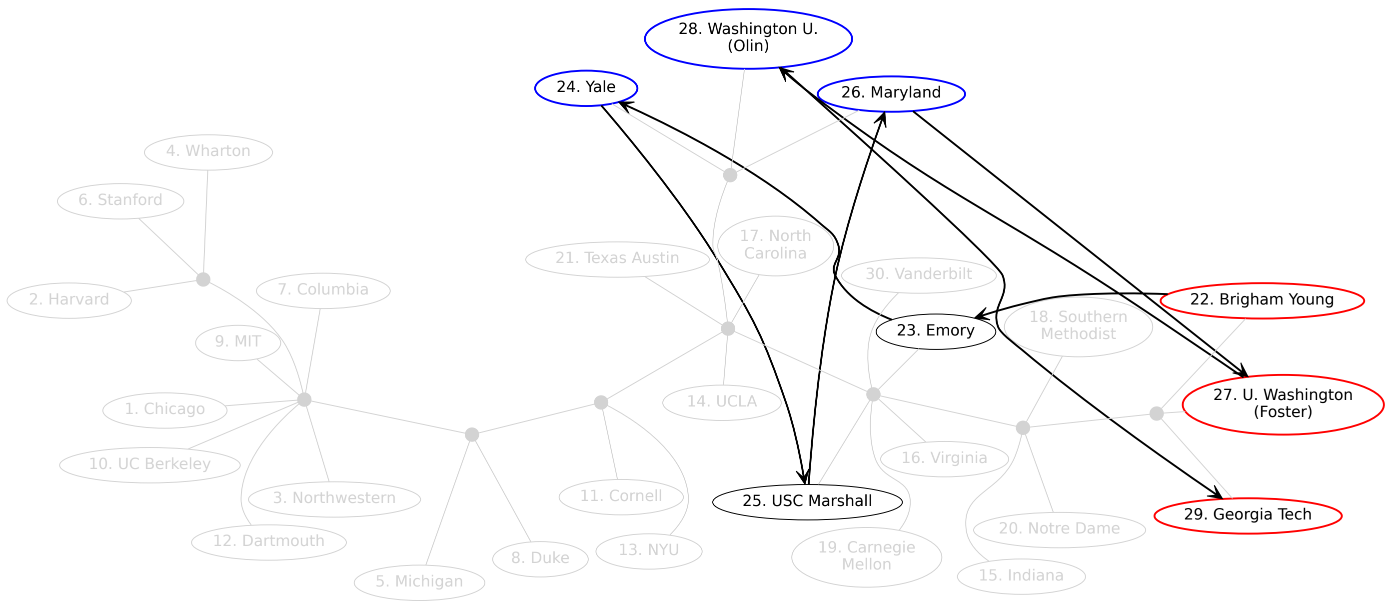


Figure 8: *Here we focus on the previously emphasized subsets. Let us introduce to graph $G = (V, E)$ a new set E^r of directed edges, in which $e^r_{i,j} \in E^r \leftrightarrow r(i) = r(j) - 1$, where $r(n)$ is the rank of school n. The graph $G^r = (V, E \cup E^r)$ plots the edges in E^r emphasized in black. Notice how the large number of edge crossings display potential rank anomalies.*

School	2008	Grad. poll	Corp. poll	Intel. capital	Tuition & fees	Post-MBA pay ($000)	Selectivity (%)	Gen. mgmt. skills	Analyt. skills
Brigham Young	22	27	15	41	$ 37,010.00	90	56	A	A
U. of Wash. (Foster)	27	30	26	29	$ 64,902.00	85	30	B	B
Georgia Tech.	29	31	28	26	$ 64,152.00	95	29	B	A
(group range)		27-31	15-28	26-41	37-64 K	85-95 K	29-56 %	A-B	A-B
Yale	24	19	33	10	$ 93,098.00	97	14	C	C
Maryland (Smith)	26	28	42	3	$ 82,435.00	91	28	C	C
Wash. U. (Olin)	28	24	41	16	$ 82,672.00	90	34	C	C
(group range)		19-28	33-42	3-16	82-93 K	90-97 K	14-34 %	C	C

Table 1: *Rank anomalies. Though schools ranked {22, 24, 26, 27, 28, and 29} seem close in the ranking, they are clearly separable into different clusters as the system discovers many dimensions in which they differ. Notice, for instance, how Brigham Young's data deviates from the the subset formed by {Yale, Maryland, Washington U.} in many dimensions.*

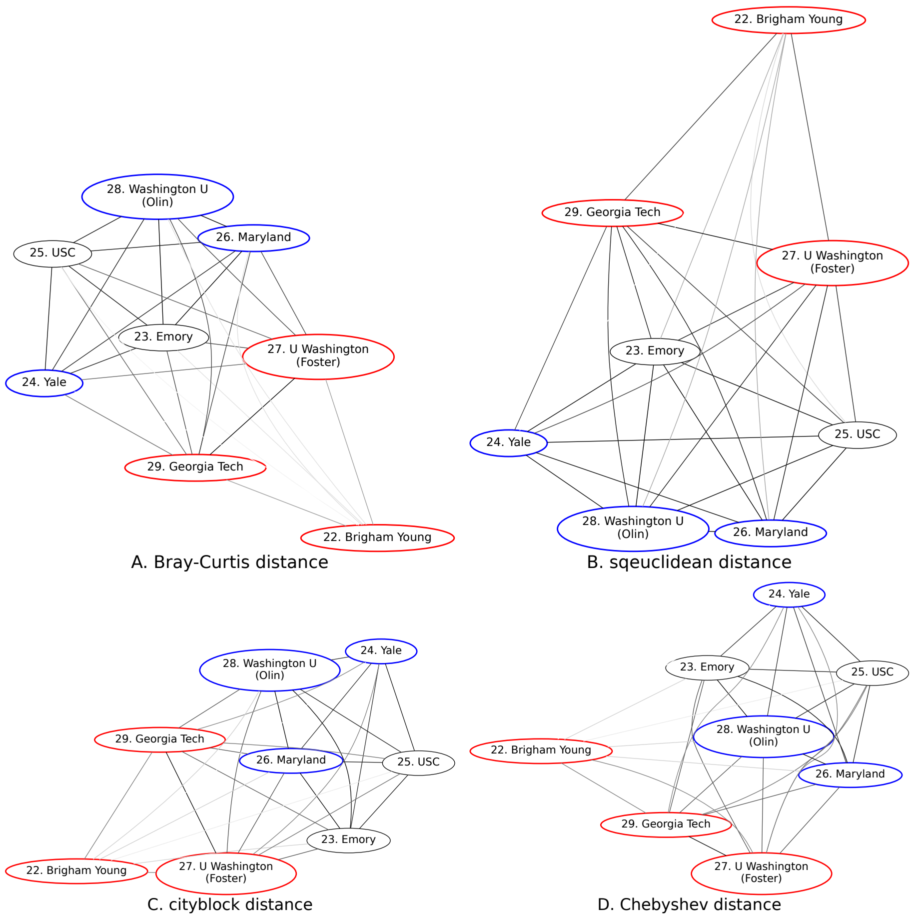


Figure 9: *What are the distances between schools? Here we see maps of distances between schools, organized under different distance metrics. Darker edge colors imply closer distances. Under these distance metrics, Brigham Young, for example, becomes increasingly "far" from Yale, and from other schools not in its subgroup.*

two different groups, obliterating their differences along the way (see Figure 8).

We find noteworthy that, to Business Week's methodology cities 'do not exist', e.g., a student living in Bloomington IN has the exact same experience of a student living in New York City—the data is simply oblivious to this information.

To sum up: if the ontology of the world of MBA programs consisted solely of the dimensions included in the rankings, if the collected data were an absolutely perfect reflection of reality, and even if the aforementioned criticisms of rankings were all invalid, this much is true: a student with a strong preference for the no. 27th school would find that school no. 29 seems a better match than school no. 24. The method proposed herein is a clear alternative to mitigate the cognitive dissonance in choosing between a school that better reflects one's true preferences versus the "better ranked" one. This type of meaningful information can be brought to light by the KT-algorithm, something rank ordering and other methods which impose structural form remain oblivious to.

5. Theoretical Contributions

5.1. Summary and open questions

We introduce, to the decision science community, Kemp and Tenenbaum's model for finding structural form in data. Instead of presenting it under the perspective of a psychological theory, our goal here is to describe it as a new methodology for research. In our experiments, we have applied the method to the data used to construct school rankings by Business Week (2008). We claim the method provides insights into the multidimensional space in which schools compete, and that the resulting KT-Structures better reflect the multi-faceted reality of business schools and are better representations than the widely disseminated rankings.

Using the very same features used in constructing the rankings, the KT-Structures bring to light anomalies in which schools may be next to each other in the ranks while bearing few resemblances in their numerous dimensions. Conversely, schools can be far in the ranks, but have a large set of similar features. We therefore question the validity of school rankings: A rank is not necessarily the most adequate form to represent (or understand) entities with no dominance relation. Additionally, statistical and data mining methods often presuppose a hidden structure, such as a cluster, a tree, or coordinates.

The MBA program rankings, furthermore, impose a representational form that is unfit for the type of information they hope to convey. This has sweeping implications towards school strategy, positioning, and, because of the wide impact of published rankings, towards prospective students and all stakeholders. One can only idealize a world in which the structures that best reflect the data are widely disseminated for public consumption.

Further investigation is warranted towards exploring the general nature of rank anomalies in different data sets, and more specifically on the nature of the space of Business School relations. We hope this introduction to the KT-algorithm is able to illuminate its highly innovative potential. It is an approach that should perhaps be admitted into our toolbox of research methods.

5.2. Theoretical contributions: Clusters, Bifurcations, and how the pseudo-science of rankings conflict with reality.

Reality is more complex than the linear projection of rankings would suggest. Some schools cluster into specializations (executive education, finance, marketing, operations management, research, etc). Some schools are more affordable, others can sustain rock-star professors and expensive programs. We find, in ranking after ranking, that Business schools bifurcate into i) clusters that share some features in stark contrast to ii) other clusters that may have the exact opposite features.

It is a simple statistical fact that, as the number of non-correlated variables grows, the probability of a bifurcation increases. Bifurcations are, in fact, inevitable. Yet, rankings compress all schools into a single orthogonal projection and, in the process, aggregate schools from completely different clusters. Rankings lose enormously valuable information about the nature of the schools. This leads us to our last proposition:

Proposition 4. *Due to their oversized influence to schools, scholars should perhaps take a step back and question the philosophy and methodology employed in rankings. Rank anomalies and bifurcations demonstrate that the one-dimensional nature of rankings should be put to question and to serious scrutiny.*

Because they use a methodology it does not follow that the institutions behind rankings produce anything of scientific value or that corresponds to the complex nuances of the real world. In fact, we accuse them of being pseudo-scientific, of having the appearance but not the substance of science. The application of a one-dimensional model to a multidimensional statistical space should not be presented to the public without seriously qualifying its limitations, its potential for error, and its minute scope. We do hope more scholars come to question the scientific validity of program rankings. Finally, we cannot refrain from noting that the due passage of time seems to have a way of dealing with scientific pretense.

6. References

Abe, M. (2009). "Counting your customers" one by one: A hierarchical bayes extension to the pareto/nbd model. *Marketing Science*, 28:541–553.

Alvesson, M. (1990). Organization: From substance to image? *Organization Studies*, 11:373–394.

Assaf, A. G. and Josiassen, A. (2012). Time-varying production efficiency in the health care foodservice industry: A bayesian method. *Journal of Business Research*, 65(5):617–625.

Assaf, A. G., Josiassen, A., and Gillen, D. (forthcoming). Measuring firm performance: Bayesian estimates with good and bad outputs. *Journal of Business Research*.

Basel, B. (2004). Families of strong kt structures in six dimensions. *Commentarii Mathematici Helvetici*, 79:317–340.

Corley, K. and Gioia, D. (2000). The rankings game: Managing business school reputation. *Corporate Reputation Review*, 3:319–333.

Cornelissen, J. and Thorpe, R. (2002). Measuring a business school's reputation: perspectives, problems, and prospects. *European Management Journal*, 20:172–178.

Dichev, I. (1999). How good are business school rankings? *Journal of Business*, 72(2):201–213.

Ehrenberg, R., Cheslock, J., and Epifantseva, J. (2001). Paying our presidents: What do trustees value? *Review of Higher Education*, 25:15–37.

Elsbach, K. and Kramer, R. (1996). Members' responses to organizational identity threats: Encountering and countering the business week rankings. *Administrative Science Quarterly*, 41:442–476.

Fee, C., Hadlock, C., and Pierce, J. (2005). Business school rankings and business school deans: A study of nonprofit governance. *Financial Management*, 34:143–166.

Florian, R. (2007). Irreproducibility of the results of the shanghai academic ranking of world universities. *Scientometrics*, 72:25–32.

Frey, B. J. and Dueck, D. (2007). Clustering by passing messages between data points. *science*, 315(5814):972–976.

Gioia, D. and Corley, K. (2002). Being good versus looking good: Business school rankings and the circean transformation from substance to image. *Academy of Management Learning & Education*, 1(1):107–120.

Gioia, D., Schultz, M., and Corley, K. (2000). Organizational identity, image and adaptive instability. *Academy of Management Review*, 25:63–81.

Hofstadter, D. and FARG (1995). *Fluid Concept and Creative Analogies: Computer Models of The Fundamental Mechanisms of Thought*. Basic Books, New York, NY.

Ioannidis, J., Patsopoulos, N., Kavvoura, F., Tatsioni, A., Evangelou, E., Kouri, I., Contopoulos-Ioannidis, D., and Liberopoulos, G. (2007). International ranking systems for universities and institutions: a critical appraisal. *BMC Medicine*, 5:30.

Kemp, C. and Tenenbaum, J. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences*, 105(31):10687–10692.

Linhares, A. (2000). A glimpse at the metaphysics of bongard problems. *Artificial Intelligence*, 121:251–270.

Linhares, A. and Brum, P. (2007). Understanding our understanding of strategic scenarios: What role do chunks play? *Cognitive Science*, 31:989–1007.

Linhares, A. and Chada, D. M. (2013). What is the nature of the mind's pattern-recognition process? *New Ideas in Psychology*, 31:108–121.

Linhares, A. and Freitas, A. (2010). Questioning chase and simon's (1973) "perception in chess": The "experience recognition" hypothesis. *New Ideas in Psychology*, 28:64–78.

Liu, N. C. and Cheng, Y. (2005). The academic ranking of world universities. *Higher Education in Europe*, 30:127–136.

Lock, K. and Gelman, A. (2010). Bayesian combination of state polls and election forecasts. *Political Analysis*, 18(3):337–348.

M. Chirgui, Z. and Penard, T. (2011). How to provide quality of service guarantees in the internet? lessons learnt from the airline and card-based payment sectors. *European Management Journal*, 29(4):306–318.

Park, H. J. and Kim, S.-H. (2013). A bayesian network approach to examining key success factors of mobile games. *Journal of Business Research*, 66(3):1353–1359.

Pfeffer, J. and Fong, C. (2004). The business school "business": Some lessons from the U.S. experience. *Journal of Management Studies*, 41(8):1501–1520.

Siemens, J., Burton, S., Jensen, T., and Mendoza, N. (2005). An examination of the relationship between research productivity in prestigious business journals and popular press business school rankings. *Journal of Business Research*, 58(4):467–476.

Smith, M., Kohn, R., and Mathur, S. K. (2000). Bayesian semiparametric regression: An exposition and application to print advertising data. *Journal of Business Research*, 49(3):229–244.

Tracy, J. and Waldfogel, J. (1997). The best business schools: a market-based approach. *Journal of Business*, 70:1–31.

Trieschmann, J., Dennis, A., Northcraft, G., and Jr., A. N. (2000). Serving multiple constituencies in business schools: M.B.A. program versus research performance. *Academy of Management Journal*, 43:1130–1141.

Van Der Merwe, R., Pitt, L., and Berthon, P. (2004). Elucidating elusive ensembles:: The strategic value of informal internet networks. *European Management Journal*, 22(1):12–26.

Zell, D. (2001). The market-driven business school: Has the pendulum swung too far? *Journal of Management Inquiry*, 10:324–338.