Dissertação:
PRIMED AND PREJUDICED: THE KURANTA-BOTHATA STUDY

Luiz Carlos Victorino de Souza Junior
Matrícula: 103105006
Agosto de 2011
PRIMED AND PREJUDICED: THE KURANTA-BOTHATA STUDY

por
Luiz Carlos Victorino de Souza Junior

Professor Orientador: Alexandre Linhares

Dissertação apresentada como pré-requisito para a conclusão do curso de Mestrado Acadêmico em Gestão Empresarial da Escola Brasileira de Administração Pública e de Empresa da Fundação Getulio Vargas.
Orientador: Professor Alexandre Linhares

Junho de 2011
Souza Junior, Luiz Carlos Victorino de
Primed and prejudice : the Kuranta-Bothata study / Luiz Carlos Victorino de
Souza Junior. - 2012.
55 f.

Dissertação (mestrado) - Escola Brasileira de Administração Pública e de
Empresas, Centro de Formação Acadêmica e Pesquisa.
Orientador: Alexandre Linhares.
Inclui bibliografia.

1. Revisão por pares. 2. Publicações científicas - Avaliação. I. Linhares,
Alexandre. II. Escola Brasileira de Administração Pública e de Empresas. Centro
de Formação Acadêmica e Pesquisa. III. Título.

CDD – 070.579
FUNDAÇÃO GETÚLIO VARGAS
ESCOLA BRASILEIRA DE ADMINISTRAÇÃO PÚBLICA E DE EMPRESAS
CURSO DE MESTRADO EM ADMINISTRAÇÃO

DISSECTAÇÃO DE MESTRADO
PRIMED AND PREJUDICE: THE KURANTA-BOTHATA STUDY

APRESENTADA POR: LUIZ CARLOS VICTORINO DE SOUZA JÚNIOR

E
APROVADA EM: 08/07/2011
PELA BANCA EXAMINADORA

Alexandre Linhares
Doutor em Computação Aplicada

Marco Túlio Fundão Zanini
Doutor em Gestão Empresarial

Jarbas dos Santos Silva
Doutor em Administração
ABSTRACT

The peer review system is a core tenet of scientific publishing, as it intends to select the studies that advance knowledge. However, there have been questions concerning biases in peer review. One example amongst many is given by the essay "Lost science in the third world" (GIBBS, 1995), which argues that there is serious bias concerning an author's address. Though there is some empirical evidence to back this statement, our knowledge of such biases, however, is limited, as few controlled studies have been provided in the literature. In this work, we have conducted an experiment, in which subjects were assigned the task of reviewing a manuscript lacking any author identification, except for a small footnote acknowledging financial support from an agency that does not exist. Under one condition, the funding agency's name is associated with the African continent, while on another condition, the agency's name is associated with the European continent. Our experiments do provide some support for the author's-bias-hypothesis.
ACKNOWLEDGEMENTS

I would like to thank God for absolutely everything – without Him, nothing of this would be possible. Also, I would like to thank my family: My parents, Luiz and Lucia, for all the support throughout the years and all the effort invested in me. My sister Livia, who fights in the academic battlefield as I do, for the example of energy and hard-work. My grandmothers Lea and Marilia, for trying to spoil me and making such a good job. Thanks to my wife Sarah for letting me build my world of dreams around her - I'm so glad that I have found you. All my family, old and new, thanks for trusting and supporting me. I hope I can make you as proud of me as I am of you.

I would like to thank Professor Alexandre Linhares for believing in my potential on my admission interview. That day changed my life. I also have learned googols of things during our meetings and this changed my life as well.

I would like to thank my friends, whom in the words of Shakespeare, are “the family we choose for ourselves”.

Finally, I would like to thank my friends from FGV: our journey was ardorous, treacherous and sometimes frightenly lonely, but we stood together and together we have succeeded. For you, I dedicate this quote from The Lord of The Rings, by J.R.R. Tolkien: “The road goes ever on and on, down from the door where it began. Now far ahead the road has gone, and I must follow, if I can, pursuing it with eager feet, until it joins some larger way where many paths and errands meet. And whither then? I cannot say”.
SUMMARY

1 INTRODUCTION ........................................................................................................ 1
2 THEORETICAL FRAMEWORK ........................................................................... 5
3 MATERIALS AND METHODS .............................................................................. 7
  3.1 Design and Procedure ...................................................................................... 7
  3.2 Participants ...................................................................................................... 7
  3.3 Materials .......................................................................................................... 7
  3.4 Data Collection ................................................................................................ 8
4 RESULTS ............................................................................................................... 9
5 CONCLUSION ...................................................................................................... Error! Bookmark not defined.
6 REFERENCES ...................................................................................................... 21
7 APENDIX A – Questionaire (Pre-test) ................................................................. 25
8 APENDIX B – Kuranta-Bothata Scientific Paper .................................................. 26
9 APENDIX C – Welgesteld-Tijdschrift Scientific Paper ........................................ 43
10 APENDIX D – Scientific Paper Evaluation Form ............................................... 44
LIST OF FIGURES

Figure 1 - After (ALLISON & LONG, 1990). The average number of publications versus time, when a researcher changes address from a high (low) prestige institution to a low (high) one. ................................................................. 3

Figure 2 - After (ALLISON & LONG, 1990). The average number of citations across time, when a researcher changes address from a high (low) prestige institution to a low (high) one. ................................................................. 3

Figure 3 - Deviation from the neutral position on the question “Do you think this paper is original?” ................................................................. 9

Figure 4 - Deviation from the neutral position on the question “Do you think the area of research of this paper will grow in the future?” ............................................. 10

Figure 5 - Deviation from the neutral position on the question “Do you think this paper is well-executed? Is the methodology sound?” .................................................. 10

Figure 6 - Deviation from the neutral position on the question “Do you think the conclusions are warranted?” ....................................................... 11

Figure 7 - Deviation from the neutral position on the question “Do you think this paper is suitable for publication?” ................................................................. 12

Figure 8 - Deviation from the neutral position on the question “Would you authorize the authors to receive funding for extending this research?” ............................................. 13

Figure 9 - Deviation from the neutral position on the question “Would you cite this paper if you eventually produce work on a related area?” ............................................. 13

Figure 10 - Deviation from the neutral position on the question “Do you think others will find this paper original?” ................................................................. 15

Figure 11 - Deviation from the neutral position on the question “Do you think others will find that the area of research of this paper will grow in the future?” ..... 16

Figure 12 - Deviation from the neutral position on the question “Do you think this paper is well-executed? Is the methodology sound?” .................................................. 16

Figure 13 - Deviation from the neutral position on the question “Do you think others will find the conclusions warranted?” ................................................................. 17

Figure 14 - Deviation from the neutral position on the question “Do you think others will find that this is suitable for publication?” ................................................................. 17
Figure 15 - Deviation from the neutral position on the question “Do you think others would authorize the authors to receive funding for extending this research?”.

........................................................................................................................................ 18

Figure 16 - Deviation from the neutral position on the question “Do you think others will cite this paper if producing work on a related area?” ....................................................................................................................................... 18

Reproduction of the Kuranta-Bothata Scientific Paper........................................26--43

Reproduction of the first page of the Welgesteld-Tijdschrift Scientific Paper..........44
1 INTRODUCTION

Science stands on the process of peer review. Peer review aims to assure reasonable and fair consideration of scientific papers—and not only for publishing in journals but also for research funding, tenure, prize awarding and so forth. Even though among journals (and institutions) there are small differences in the procedure itself, the process is based on the concept of scientific experts analyzing and scrutinizing studies in their specific field, in order to determine which ones are publishable, how revisions should be conducted, or if publication should be supported by a journal.

Such evaluation is carried out to verify if a study is well conducted, its methodology is adequate and, obviously, if the results are significant. Since those specialists conducting the review are experienced in publishing in the field, it is reasonable to assume that the process enhances the development of specific knowledge.

Nevertheless, the equality and transparency of the process is far from achieving a consensus, and questions concerning biases in the traditional peer review have been brought to attention in many fields. Many journals are currently experimenting with alternative models, such as: i) double-blind review, in which not only referees are undisclosed, but also the names and addresses of authors; ii) published review, in which the review (and referee's name) is made public, alongside the published article; or even iii) open peer review, in which submitted papers can be reviewed by anyone with an interest in reviewing a particular paper (for a well-known experiment in this regard conducted by the journal Nature, see Greaves et. al., 2006).

A number of criticisms have been laid out in the literature, particularly concerning bias against author's addresses (sometimes low-prestige institutions from high-income countries, sometimes institutions from low-income countries). One such example is the essay “Lost science in the third world” (Gibbs, 1995) which argues that there is a bias concerning an author’s nationality. Gibbs argues, for instance, that China and India, accounting for more than 45% of the world's population, published barely 3% of papers in the ISI database. (Of course, there are two qualifications here: the percentage of educated scientists in low-income countries is much smaller than in high-income countries; moreover, this is data from the 1990's; numerous global changes have occurred since.) Gibbs also points out claims of bias against authors from low-income countries:
"There is no question in my mind that there is some inherent prejudice in the minds of some referees in the West about authors from Third World countries," says C.N.R. Rao, president of the Jawaharial Nehru Center for Advanced Scientific Research in Bangalore. "Referees tend to feel that good work could not have been done in a developing country. I have published in major journals of the world for the past 40 years, but even now I face some prejudice from referees. This is not infrequent." (GIBBS 1995, p. 82)

Some scientists who have moved from industrial to less developed countries have found that the address change makes a difference to reviewers.

"When I was a resident in Boston, I was able to publish papers in the American Journal of Pathology with a couple of well-known pathologists," Benítez says, "They flew through to publication with no problem. After that, I went to the University of Bonn in Germany and published two papers in Nature. Then I came back to Mexico with more experience and maturity. But now when I have sent papers to the same journals, they have been rejected immediately" (GIBBS 1995, p.81--82)

Other studies have also indicated flaws on the peer review process, such as a paper published in Science, which have found that getting a research grant, even though conducted through a peer review process, depended "to a significant extent on chance" (Simon, Cole & Cole, 1981, p.881).

More recently, when investigating publication in top journals, researchers have found biases in the peer review process concerning academic origin and affiliation (Long, Bowers, Barnett & White, 1998; Meneghini, Packer & Nassi-Calò, 2008). In the USA, a study by Allison & Long (1990) compares the number of papers published, and citations received yearly, across the years, by authors that made a career move. They find that researchers that change positions to lower-prestige institutions within the U.S. publish less and receive fewer citations after the career move. On the other hand, those that move up the prestige hierarchy (i.e., change jobs to a top-caliber department), see their numbers of papers and citations increasing after the job change. It is plausible that
factors other than competence and maturity influence scientific recognition, and our aim in this work is to conduct a controlled experiment of the peer review process.

Figure 1. (After ALLISON & LONG, 1990). The average number of publications versus time, when a researcher changes address from a high (low) prestige institution to a low (high) one.

Figure 2. After (ALLISON & LONG, 1990). The average number of citations across time, when a researcher changes address from a high (low) prestige institution to a low (high) one.
Despite the fact that there is some empirical evidence to back the idea of biases in the peer review process, our knowledge of such biases is limited, as few studies have been provided in the literature, and most of them concern specificities of different fields, and not the phenomenon itself.

As the basis of the peer review process is the evaluation and judgment of specialists (CECI & PETERS, 1982), to understand and avoid the biases in the peer review process, it is necessary to investigate how paper evaluation (and judgment concerning the publish-ability of a paper) can be affected, hence, influencing the peer review process (SMITH, 2006). The next chapter provides a background of the psychological mechanisms of prejudice and stereotyping: priming.

Our objective is to investigate if through a priming effect, subjects will show a bias in the evaluation of a scientific paper.
2 THEORETICAL FRAMEWORK

There are numerous situations in which stereotyping and judgment are connected, and understanding this relation can improve evaluation processes (Herr, 1986; Steele and Aronson, 1995; Stangor, 2000). On the present study, we focus on scientific papers evaluation, and how the stereotype of a low-prestige institution, or of a low-prestige country, can affect evaluation and judgment of a scientific paper. There is literature showing stereotyping in judgment and evaluation of scientific papers (Swim, et.al., 1989; Monaham, Brown & Shtrulıs, 2003) and, since the works of Cole and Cole (1973) and Hagstrom (1968) it seems that scientists at prestigious universities tend to have higher rates of publication and higher rates of citation, in part attributable to the prestige of their institutions--even if we consider that prestigious institutions have better research centers, superior infrastructure, attract better scientists and that, developed countries usually have well-developed scientific incentive programs and higher investments on science, it is reasonable to think that this will have an impact on their scientific production.

It is important to highlight that stereotyping and prejudice are different constructs. Steele and Aronson (1995) described stereotypes as “standardized and simplified conceptions of groups based on some prior assumptions”. Actually, stereotyping is a cognitive strategy, with evolutionary origins stemming from the beginnings of the human race, to optimize mental processing. After labeling a person, an object, or a situation, it is easier for the brain to identify and react (Kleg, 1993). Researchers have even shown that, depending on the situation, stereotypes could be very accurate (Nelson, 2009; Lee, Jussim & McCauley, 2005). The interesting fact about stereotyping, and the focus of the present study, is that stereotypes can be activated even when the stimulus presented is unknown--that is, when the brain makes connections with past experiences in order to develop a subsequent coherent response (Macrae, Stangor & Milne, 1994; Oliner, 2001). Which means that, even when something is shown for the first time, based on our previous experiences, we automatically develop a labeling process which will influence our present and future responses to that stimulus (Greenwald & Banaji, 1995).

Stereotypes arise from unconscious associations, and numerous recent studies in cognitive psychology have probed the mechanisms with which we activate this
associations unconsciously and effortlessly. A particularly intriguing cognitive phenomenon known as priming has been gaining more attention in the last years, not only for its potential to influence behaviour, but also for its widespread presence (Bargh, Brownell & Harris, 2009). Priming could be defined as “a process in which an individual is exposed to a stimulus that influences the response to a subsequent stimulus” (Kolb & Wishaw, 2003, p. 453). The literature on priming is extensive on how it can influence the basic processes of evaluation and judgment (Phillipot et al. 1991, Mussweiler & Strack, 2000), and the consequences of this influence in different scenarios, for example, in financial behavior (Gilad & Kliner, 2008).

The cognitive process behind priming is still not completely described, but the standard perspective is that, once a particular stimulus is shown, it activates neural networks through a process called spreading activation (Reisberg, 2007)--the stimulus activates pieces of information in memory that are related or associated to the content, influencing the response to the stimulus (Bargh & Chartrand, 2009). And since the activation of those networks is not conscious, there is no participation of conscious will in the process (Bargh, 1982).

A good example is the research conducted with rude and polite stimuli, when subjects primed with rude words were significantly more likely to interrupt a person than those primed with neutral words. And also, the subjects primed with polite words were the least likely to interrupt in the same situations (Bargh, Chen & Burrows, 1996). Even a simple stimulus, like holding a hot or cold beverage just before an interview has influenced the opinion of the subjects towards the interviewer (Williams & Bargh, 2008). It is well accepted – and different researchers have been proving - that not only simple and basic responses can be primed, but even extremely complex cognitive processes can be affected nonconsciously through priming, such as use of social norms (Aarts & Dijksterhuis, 2003), egalitarian behavior (Bargh, et al 2001), cultural ideologies (Nisbett, 2003) and, the focus of the present study, judgment and stereotypes (Bargh, 2006; Ferguson, Bargh & Nayak, 2005; and Dijksterhuis & Bargh, 2001).

We hypothesize that a scientific paper, if presented as funded by an European-sounding institution, will be better evaluated than the same paper, if presented by an African-sounding institution.
3 MATERIALS AND METHODS

3.1 Design and Procedure

Participants were given a scientific article to read with no time restraint. After reading the article and giving it back, they should answer a questionnaire to evaluate the article.

On the article given, a footnote contained a piece of information to trigger the priming effect. A pretest was conducted to evaluate the effect of the priming factors used on the main experiment.

3.2 Participants

For the pretest, 63 undergraduate students of business administration from a major Brazilian university participated voluntarily.

For the main experiment, participants were randomly selected from the largest administration academic conference in Brazil, in a total of 80 participants between coordinators, professors and students from master and doctor programs, who also participated in the experiment voluntarily.

3.3 Materials

A seventeen-page scientific article on the area of general management was formatted as such, but with the only information available being the text itself and the footnote regarding the funding institution. There was no other information disclosed, such as author’s name or university.

There were two conditions: On condition 1, the footnote informed that the research was financed by the High Scientific Council of the Welgesteld-Tijdschrift Province. On condition 2, with the exact same paper, the footnote informed that the research was financed by the High Scientific Council of the Kuranta-Bothata Province.

Both institutions are not real and neither are the provinces. Actually, Welgesteld-Tijdschrift means “wealthy magazine” in Dutch, and Kuranta-Bothata means “problematic magazine” in Setswana, a Southern African language, spoken in Botswana, Zimbabwe and Namibia. Both languages are not well known in Brazil, and a
pretest was conducted to test the hypothesis that, for Portuguese speakers, the word Welgesteld-Tijdschrift resembles some sort of European language, and, on the other side, Kuranta-Bothata has an african similitude. The results of the pretest confirmed this hypothesis.

3.4 Data Collection

For the pre-test, a one-page questionnaire was used to determine the priming effect of the words chosen for the funding institutions (“Welgesteld-Tijdschrift” and “Kuranta-Bothata”). This questionnaire had five questions. The first four, asking which of those institutions were more prestigious, to which would the respondent apply for a master program, which seemed more reliable as far as scientific production was concerned and from which institution would they like to receive information about courses, post-graduate programs and academic exchange. On the last question, respondents were asked about the origin country of each institution and which of those seemed more social and economic developed.

For the main experiment, a three-page questionnaire was used. It was divided in three parts: The first part had questions regarding the quality of the paper, the methodology, the probability of publication, and if the research should receive new funding. The second part had the same questions, but the subjects should answer how other scientists would evaluate the paper. The third part of the questionnaire asked if they remembered any information about the funding institution and which aspects were more relevant to undergird their opinions.
4 RESULTS

The questionnaires were analyzed based on the responses given through a Likert Scale, from 1 (completely disagree) to 5 (completely agree). Given that 3 was the neutral response (not agreeing or disagreeing), for a better visualizations of the results, we have translated the scale to a scoreboard (e.g., the scale's neutral response 3 is equal to 0, and the scale 1 to 5 was translated to the range [-2,+2], and we look at the sum totals).

![Figure 3. Deviation from the neutral position on the question “Do you think this paper is original?”](image)

On this item the results were not meaningful. The Kuranta-Bothata (KB) condition got a score of 15, while the Welgesteld-Tijdschrift (WT) condition got a score of 19. Nevertheless, this question is tangential to the effect, since originality it not necessarily correlated with quality--a paper can be original but not well structured or poorly conducted, and vice-versa.
Similarly to the previous item, this question showed no meaningful difference between the two conditions, with KB reaching 15 points and while WT reaching 18, and again, this question was tangential to the effect because the growth of an area is not related to the quality of a paper. There are good papers in declining areas and low-quality papers in growing areas.

Figure 5. Deviation from the neutral position on the question “Do you think this paper is well-executed? Is the methodology sound?”.
Even though the scores for both conditions were positive, a meaningful difference was found, with the KB condition reaching 1 point and the WT condition reaching 16. That is, subjects seem neutral in regards to KB, but have a generally positive view of the WT condition. Unlike the first two questions, the execution and the methodology of a paper are linked to its quality, which could indicate an effect caused by the conditions presented.

![Bar Chart](image)

**Figure 6.** Deviation from the neutral position on the question “Do you think the conclusions are warranted?”.

On this item, the results indicated a negative tendency on the KB condition, which scored -8 against +9 from the WT condition. Even though a paper can be well executed but reach no warranted conclusions, a negative score means that most of the respondents on the KB condition perceived signs of problems with the paper’s conclusions. This result endorses the idea of the conditions affecting the perception of the quality of the paper.
Figure 7. Deviation from the neutral position on the question “Do you think this paper is suitable for publication?”.

For this question, the KB condition reached -6 points while the WT condition reached 19. Not only the negative tendency for the KB condition, but also the positive and expressive score of the WT condition, reinforce the effects of priming in the perception of quality.

Being suitable for publication is a good indicative of a quality of a paper and even when the editorial differences are taken into account, a scientific paper has to reach a standard of quality to be published, which makes this question extremely relevant to the investigated effect.
Receiving funds for extending the research not only is related to the quality of a paper, but also indicates that it has potential to reach relevant results in the future. The results for this item were -13 for the KB condition and +20 for the WT condition, which gives a good evidence that the priming stimulus had an effect on the perception of quality.
With a score of -15 for the KB condition and +21 for the WT condition, this item showed the largest difference between the results achieved. Not only it indicates the negative effect caused by the KB condition and the positive effect caused by the WT condition, but also this is the only item which contains a more personal/behavioral aspect. The act of citing a scientific paper not only denotes the quality of the cited paper, but through linking it to the work on progress, also indicates the relevance of the contribution given by the cited article.

Analyzing the results for the first part of the questionnaire (in which the respondents showed their own opinions about the paper), the effect of the priming stimulus on the subjects was evidenced by the negative scores reached by the KB condition and the positive scores reached by the WT condition. The respondents on the KB condition showed a tendency of giving negative grades for the paper in most of the evaluated aspects, and similarly, the respondents on the WT conditions evaluated the paper more positively.

Not only the negative and positive trends are relevant, but also the quantitative differences in the scores for both conditions—perceptible mainly on the aspects directly linked to the quality of the paper.

The tangential aspects showing a lighter effect, actually ends up reinforcing the evidence on the priming effect causing the perception differences between KB and WT conditions, since the more linked to the quality the aspect was, the larger the difference between the scores.

Evidently, it does not mean that the peer review system has a established bias, and that is not even the objective of the present study, but it indicates that a subtly presented stimulus (such as two words on a footnote) can trigger a priming process that alters the perception of the subject, hence affecting its judgment towards an object.

If we consider that Kuranta-Bothata and Welgesteld-Tijdschrift are not even real as research funding institutions, it indicates that not only the subjects were affected by the priming stimulus, but also that they were able to link those words, through spreading activation, to contents that already existed in their cognition—as shown both in the pre-test and in the main experiment. The sounding-like effect of the words “Kuranta-Bothata” and “Welgesteld-Tijdschrift” was sufficient to activate pre-existing concepts, consequently affecting the perception and judgment of the paper.
Even though the subjects were not all reviewers, all of them had experience in evaluating papers, since it is part of their daily activities, as either graduate students or faculty.

On the second part of the questionnaire, the respondents were asked how others would evaluate the scientific paper in the same aspects as they had evaluated it. The objective of this part was to investigate if the conditions would be present not only on the respondents’ opinion, but also on their impressions on the opinions of others.

The results on the second part, shown on figures 10--16, followed the same trends perceptible on the first part, but were less striking. Although we were unable to identify the causes for this effect, we hypothesize that the respondents were uncertain or unsure of the opinion of others, and tended to score more neutrally.

![Figure 10. Deviation from the neutral position on the question “Do you think others will find this paper original?”](image-url)
Figure 11. Deviation from the neutral position on the question “Do you think others will find that the area of research of this paper will grow in the future?”.

Figure 12. Deviation from the neutral position on the question “Do you think this paper is well-executed? Is the methodology sound?”.
Figure 13. Deviation from the neutral position on the question “Do you think others will find the conclusions warranted?”.

Figure 14. Deviation from the neutral position on the question “Do you think others will find that this is suitable for publication?”.
Figure 15. Deviation from the neutral position on the question “Do you think others would authorize the authors to receive funding for extending this research?”.

Figure 16. Deviation from the neutral position on the question “Do you think others will cite this paper if producing work on a related area?”.
5 CONCLUSION

Numerous authors have argued that there may be bias in the peer-review system (Langfeldt, 2006; Marsh, Jayasinghe & Bond, 2008; Lane & Linden, 2009). In the aforementioned "lost science in the third world" study, for instance, some Brazilians (among others) argue that there is serious bias against authors from low-income countries (including Brazil). In this work we turn the tables and argue that Brazilians also display bias towards low-income countries--our controlled experiment shows subjects heavily skewed towards authors from European-sounding institutions, as opposed to authors from African-sounding institutions.

Perhaps our results can be generalized to the pool of scientific journal referees. But perhaps that would be a stretch too far of what we have found here. Certainly, our evidence is limited and our experiment does not immediately warrant this conclusion, for two reasons. First, consider the timing and effort involved in reviewing a journal paper. Referees, depending on the specialty and the journal, may have from weeks to months to evaluate a manuscript. Of course, in practice, referees do not use this allotted time exclusively on manuscript evaluation, but must balance this task against a background of other activities.

It is also possible that subjects that usually referee articles would reject our request for participation in this study, lacking the incentive of formal journal recognition to carry out a task that they may find routine and perhaps burdensome. It is hard to control against this adverse selection effect, unless one selects as subjects only those with journal refereeing experience; this is something we believe should be addressed in further work.

However, even if our subjects may not be a good reflection of the referee pool, and even if we may have a particularly inadequate sample in what refers to top-quality business and management journals (e.g., few of our subjects were authors of papers in such journals), all of our subjects could have been referees for the Enanpad conference (which is the world's 2nd largest management conference by number of presentations), and there is nothing to suggest that this sample of subjects is specially biased or prejudiced.

Herein lies the serious question brought by this study. We do not claim to have
shown that there is an inherent bias in the peer review system. What we claim is that the change of a mere two words, if these words bind the authors to some plausible address (an indirect measure of prestige), then the mere presence of such words can alter perception of the quality of an entire 17-page paper. The priming of an African-sounding name has an effect on perception of quality, as compared to the priming of an European-sounding name. This is the effect we have found; and we have no explanation other than prejudice to account for the disparities in the paper-quality evaluations. We do believe that further work must address, in different and perhaps larger populations, and by journal editors, this most crucial question lying at the heart of the scientific process; this question of "primed and prejudice".
REFERENCES


7 APENDIX A – Questionnaire (Pre-test)

Questionnaire

Welgesteld-Tijdschrift and Kuranta-Bothata are two foreign institutions that are planning to start working with Brazilian students. Since both institutions are not known in Brazil, it is important to evaluate the perception of Brazilian students regarding the institutions. Please, answer the following questions simply choosing one of the institutions.

Which institution is more prestigious?

(  ) Welgesteld-Tijdschrift (  ) Kuranta-Bothata

For which institution would you apply for a master program?

(  ) Welgesteld-Tijdschrift (  ) Kuranta-Bothata

Which institution seems to be more reliable, as far as scientific production is concerned?

(  ) Welgesteld-Tijdschrift (  ) Kuranta-Bothata

From which institution would you like to receive information on courses, postgraduate programs and academic exchange opportunities?

(  ) Welgesteld-Tijdschrift (  ) Kuranta-Bothata

From which country or continent do you believe these institutions are from?

Welgesteld-Tijdschrift ___________________________

Kuranta-Bothata ___________________________

Which of those institutions seems to be from a more social and economic developed country?

(  ) Welgesteld-Tijdschrift (  ) Kuranta-Bothata

Thank you.
The multidimensional road to Harvard; or why ranking schools does not make sense

Abstract

Most quantitative research methods impose an a priori structure on the results. However, in a large number of cases, it may be of interest to researchers to discover what type of structure seems to best represent the collected data. The purpose of this paper is to present a new statistical method, KT-structures, and demonstrate i) how it is readily applicable for management scholars and ii) how it can provide insights not found on a priori-imposed structures. As a first application of KT-structures in managerial settings, we show how to use the method on the Business Week ranking of top US-based MBA programs. The resulting analysis shows that the structures imposed by ranks (orders) cannot capture the multidimensionality of the space in which Business Schools compete—even when restricting ourselves to the very few (12) dimensions used by Business Week. We place our analysis on the larger literature of critiques of School Rankings. Finally, we provide a tutorial on how researchers can take advantage of this new model. We hope readers will receive this introduction to KT-Structures with the recognition that this is a promising innovative approach that deserves to be admitted into our toolbox of research methods.

Key words: KT-Structures, Exploratory Statistical Analysis, Business School Rankings, Massively Multidimensional Spaces, Computational Cognitive Science, Bayesian Reasoning

1 Introduction

Business schools are regularly ranked by Business Week, The Economist, US News & World Report, Fortune, Financial Times, the Wall Street Journal, amongst many other organizations and periodicals. A rank is a mathematical structure also known as an order: given two distinct entities $e_1$ and $e_2$, the statement $e_1 \prec e_2$ denotes that $e_1$ precedes $e_2$. The stated meaning in a school ranking is that if school $e_1$ precedes school $e_2$, then, generally, $e_1$ should be preferred to $e_2$ by prospective students, by faculty in search of job positions,

* This work has been generously supported by grant 26/110.790/2009 from the High Scientific Council of the Kuranta-Bothata Province.

Preprint submitted to Elsevier

September 21, 2010
by potential employers of alumni, and by other observers and stakeholders. 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, for the sake of argument, the imaginary land of Simplicia. In Simplicia, there are four business schools. Two business schools, \( LE \) and \( LA \), are found at the island of Laputa—to borrow from Jonathan Swift—and are deeply concerned with theoretical development and (quite literally) 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 schools \( LE \) and \( LA \), though: \( LE \) 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 \( RE \), are found in the land of Recordia—a land in which everything must be recorded. These schools 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. In Recordia, philosophy, mathematics, statistics, and metaphors have been banned. At the start of the school year, a lottery selects one thousand examples to be taught that year, with no logical sequence between them. As in Laputa, the only difference between the schools is that \( RE \) is expensive and \( RA \) is affordable.

![Diagram of Simplicia's schools]

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

What is the structure that relates the schools of Simplicia? There are at least two equally plausible structures: a grid, or a ring. Note that, in this simplest of examples, the ring and grid will have the same characteristics— but if we add another school or 'country' their properties will then differentiate.

- 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 (ob-
advance from cognitive scientists Charles Kemp and Joshua Tenenbaum (2008) has enabled the *automatic discovery of form*. While Kemp and Tenenbaum are mostly interested in 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.
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 & Waldied 1997) or better.
clustering method to a set of data, one is assuming that clusters provide a suitable form to analyse 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 analyse the data provided or to lead prospective students to the optimal decision concerning a choice of schools.

There are two important ideas involved in their model: i) the use of a hierar-
\( P(S|F) \) is given by the number of structures compatible with a given form:

\[
P(S|F) \propto \begin{cases} 
\theta^S & \text{, i.e., if } S \text{ is incompatible with } F, \text{ then } P(S|F) = 0. \\
0 & \text{otherwise}
\end{cases}
\]

otherwise it can be computed given additional info, such as the Stirling number of the second kind and the number of \( k \)-cluster structures for a given form, as described in Kemp and Tenenbaum (2008). Graphs with numerous clusters are penalized through parameter \( \theta \).

We refer the interested reader to Kemp and Tenenbaum 2008 for the mathematical details of \( P(S|F) \) and for \( P(D|S) \), the probability of structure \( S \) given data, respectively.
cluster with all objects contained therein. This enables the 'splitting' process to continue until a final structure is reached. 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.

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

3 The BusinessWeek 2008 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-Structure. We computed all the 21 possible forms provided in the method, and
compute the structures based only on the underlying data, and does not need to understand the meaning imbued in each dimension.

3.2 Numerical experiments

The most interesting computed form is a hierarchy (e.g., a type of tree); one of which is presented in Figure 2. At a macro level, this form has some semblance with the original ranking (Figure 3). The KT-Structure distance between two schools \((i, j)\) is measured by counting the number of edges from the origin
**Figure 2.** The generated KT-Structure: an undirected hierarchy with no self-links.
Yale will charge $93098. 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).

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.

Of course, by choosing Yale over Georgia Tech, there are also significant gains—moving, however, 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, while 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%.
according to rank, though we have no way to test this at this point.) This is the type of meaningful information which the KT-Structure brings to light and which a simple rank ordering remains oblivious to.

Of course, analysis of the KT-Structure may also be valuable to faculty. When a school decides to work on improving one of its many variables, it is trying to break away from the current cluster and slowly move upstream. The KT-Structure enables a comparison to other schools in the same cluster and, moreover, highlights the differences between the clusters upstream. A school can move faster if it knows exactly where it is located in this multidimensional space, and sensitivity analysis can be conducted through small variations of parameters. Rather like Nature, Business Schools non factum salutum. There are no sudden jumps here; as there are many multidimensional curves on the road to Harvard.

4 Summary

We introduce, to the organization science community, Kemp and Tenenbaum’s model for finding structure 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 a business school education 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 the anomaly that 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. Statistical and data mining methods often presuppose a hidden structure, such as a cluster, a tree, or a ranking. The MBA program rankings, however, impose a representational form that is unfit for the type of information they hope to convey. This has sweeping implications to school

---

2 The current recession notwithstanding.
3 We use Harvard here not as an endorsement or any other judgement of value. Because of its great wealth, history, faculty, alumni, and many honors, it can be argued that Harvard University—and HBS—has become the "stereotypical world-class" University.
strategy, positioning, and, because of the wide impact of published rankings, for 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.

Appendix. Applying KT-Structures: a tutorial for social scientists

In this appendix we provide a step-by-step tutorial, so that other researchers may promptly apply their own datasets to this new method.

On the information-processing of the method

The method works by pressuposing initially that all entities (in our case, schools) are contained in a single cluster. The method then, given a specified form F and the dataset D, searches for the best structure that represents the data. In the online supplement we present a video of the method’s convergence, from a single all-encompassing cluster to a series of ‘splits’ and re-adjustments.

Software Requirements

Kemp and Tenenbaum host their code and data sets at charleskemp.com/code/formdiscovery1.0.tar.gz. The code is written in the Matlab (Matrix Laboratory) framework, which is a proprietary software, though widely available. We are at this stage attempting to execute the code in the open-source alternative, GNU Octave (www.gnu.org/software/octave/). We are also starting a translation to JAVA. The code also has dependencies on the open-source GraphViz package (graphviz.org/), an advanced package that enables numerous functions for drawing all kinds of graphs and trees.

Tutorial

There are many steps that need to be taken in order to execute the method in a new dataset. The following files must be configured:

File `setps.m`: This is one of the parameter configuration files. It has the vectors

```matlab
ps.data = {'demo_chain_feat', 'demo_ring_feat', 'demo_tree_feat', 'demo_ring_ral_bin',
'demo_hierarch_ral_bin', 'demo_order_ral_freq', 'synthpartition', 'synthchain', 'synthring'};
```
'synthtree', 'synthgrid', 'animals',... 'judges', 'colors', 'faces', 'cities',...
'mangabeys', 'bushcabinet', 'knurling',... 'prisoners', 'schools');

and

ps.dics = {{b, 'demo_chain_freq'], [b, 'demo_ring_freq'],... 
[b, 'demo_tree_freq'], [b, 'demo_ring_rel_bin'],... 
[b, 'demo_hierarchy_rel_bin'], [b, 'demo_order_rel_freq'],... 
[b, 'synthpartition'], [b, 'synthchain'], [b, 'synthring'],... 
[b, 'synthtree'], [b, 'synthgrid'], [b, 'animals'],... 
[b, 'judges'], [b, 'colors'], [b, 'faces'], [b, 'cities'],... 
[b, 'mangabeys'], [b, 'bushcabinet'], [b, 'knurling'],... [b, 'prisoners'], [b, 'schools']);

The user must include the name of the new data file in both these vectors. In our case, the inclusion in ps.data and in ps.dics is of the last entry, 'schools', and also [b, 'schools'], correspondingly.

File setrunps.m: This file needs to be altered according to the nature of the data. Is it feature data? Is it similarity data? Is it relational data?

File master.run.m: this is the main program file. A number of small changes must be made here. First, the directory path in which GraphViz is installed must be set. For example, in a windows machine:

[s,w] = system('C:\Program Files\Graphviz\bin\gvedit.exe');

The following vectors also need to be changed:

thisstruct = [1,3,6];

and

thisdata = [1:5];

Note that the numbers here must reflect the positions given in the vector ps.structures and ps.data (both found on file setups.m). In the above example, the system will load the first five entries of ps.data, each at a time, as input to search for structures, and it will search for the types of structures in entries 1, 3, and 6 of ps.structures (which are 'partition', 'order', and 'tree').

Obviously, the data files must be in the \data subdirectory.

After these steps are complete, typing master.run at the Matlab prompt will
start the execution the program. We hope readers will receive this introduction to KT-Structures with the recognition that this is a promising innovative approach that deserves to be admitted into our toolbox of research methods.

References


The multidimensional road to Harvard: or why
9 APENDIX C – Welgesteld-Tijdschrift Scientific Paper (the other 16 pages are the same as on Apendix B).

10 APENDIX D – Scientific Paper Evaluation Form

**Questionnaire**

**Paper Evaluation Form**

Given the time you have allocated to this paper, how would you rate it in the following 7 dimensions?
#1. Do you think this paper is original?
1. Not at all original
2. 
3. 
4. 
5. Extremely innovative

#2. Do you think this paper is well-executed? Is the methodology sound?
1. There are serious flaws
2. 
3. 
4. 
5. Extremely well-executed

#3. Do you think the conclusions are warranted?
1. The conclusions do NOT follow from the paper
2. 
3. 
4. 
5. The conclusions clearly follow from the paper.

#4. Do you think the area of research of this paper will grow in the future?
1. Not at all growth
2. 
3. 
4. 
5. This area will have immense growth

#5. Would you authorize the authors to receive funding for extending this research?
1. Not at all funding
2. 
3. 
4. 
5. Absolutely, further research deserves funding

#6. Do you think this is suitable for publication?
1. Not at all this paper should be published in a top-journal
2. 
3. 
4. 
5. this paper should be published in a top-journal

#7. Would you cite this paper if you eventually produce work on a related area?
1. Not at all 
2. 
3. 
4. 
5. Yes, absolutely.

Given the time you have allocated to this paper, how would do you think OTHERS would rate it in the following 7 dimensions?

#8. Do you think others will find this paper original?
1. Not at all original
2. 
3. 
4. 
5. Extremely innovative

#9. Do you think others will find this paper well-executed? Would others find the methodology sound?
1. There are serious flaws
2. 
3. 
4. 
5. Extremely well-executed

#10. Do you think others will find the conclusions warranted?
1. The conclusions do NOT follow from the paper
2. 
3. 
4. 
5. The conclusions clearly follow from the paper.

#11. Do you think others will find that the area of research of this paper will grow in the future?
1. 2. 3. 4. 5.
Not at all This area will have immense growth

#12. Do you think others would authorize the authors to receive funding for extending this research?
1. 2. 3. 4. 5.
Not at all Absolutely, further research deserves funding

#13. Do you think others will find that this is suitable for publication?
1. 2. 3. 4. 5.
Not at all This paper should be published in a top-journal

#14. Do you think others will cite this paper if producing work on a related area?
1. 2. 3. 4. 5.
Not at all Yes, absolutely.

=================================
Thank you for participating in this study!

Can you guess what the intent of this research is?
NO YES:

Would you like to receive a copy of the paper you have evaluated?
NO YES, my email is

Would you like to be informed of the intent and results of this study?
NO YES, my email is

What is your highest degree? ________________
PQ/CNPq? ________________

What is your affiliation?

What is your position?

Male [] Female []