

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE RELAÇÕES INTERNACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM RELAÇÕES
INTERNACIONAIS

JUSTIN HUGO AXEL-BERG

**COMPETING ON THE WORLD STAGE: the Universidade
de São Paulo and Global Universities Rankings.**

São Paulo
2015

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE RELAÇÕES INTERNACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM RELAÇÕES
INTERNACIONAIS

**COMPETING ON THE WORLD STAGE: the Universidade
de São Paulo and Global Universities Rankings.**

JUSTIN HUGO AXEL-BERG

Dissertação apresentada ao Programa de
Pós-Graduação em Relações
Internacionais do Instituto de Relações
Internacionais da Universidade de São
Paulo, para a obtenção do título de
Mestre em Ciências – Programa de Pós-
Graduação em Relações em
Internacionais

Orientador: Prof (Dr) Jacques Marcovitch

Versão corrigida / A versão original se encontra disponível na Biblioteca do
Instituto de Relações Internacionais e na Biblioteca Digital de Teses e
Dissertações da USP, documentos impresso e eletrônico.

São Paulo
2015

ABSTRACTS

As higher education becomes ever more universalised, and an increasingly fundamental driver in development of both human and technological capital, the need to assess and compare these capabilities has become an important feature of 21st century academia. Public debate over universities has crystallised around universities rankings. This dissertation explains the rise of the three main global university rankings and their relevance to universities on the periphery, in this case the Universidade de São Paulo. By looking first at their theoretical context and then at their methodological limitations, the dissertation shows that these rankings are inadequate measures of performance for modern research universities, and do not provide realistic or desirable goals for universities outside of the Anglo-American system. However, they do point to a variety of weaknesses within USP's system of evaluation, which are addressed with suggestions on how to improve evaluation. The final section builds on the incipient benchmarking movement to suggest a variety of different ways USP can engage with the need to compare and evaluate without the pitfalls of universal rankings.

Enquanto o ensino superior torna-se cada vez mais universalizado, e um fator determinante em desenvolvimento, tanto humano, quanto tecnológico, a necessidade de avaliar e comparar capacidades tornou-se um aspecto chave da academia no século XXI. O debate público sobre as universidades cristalizou-se no tema dos rankings. Esta dissertação busca explicar a predominância dos três maiores rankings globais, e sua relevância para as universidades de países emergentes, neste caso a Universidade de São Paulo. Analisando primeiro o seu contexto teórico e depois as suas limitações metodológicas, o trabalho revela que esses rankings são medidas inadequadas de desempenho para instituições modernas de pesquisa, e não fornecem nem metas realistas nem desejáveis para universidades fora do sistema Anglo-Americano. Todavia, os rankings apontam as várias ineficiências no sistema atualmente utilizado pela USP, e a necessidade de mudanças. A parte final propõe, com base num movimento incipiente de “*benchmarking*”, alternativas a serem exploradas pela USP tendo em vista um maior engajamento na comparação internacional para sua avaliação, sem as armadilhas inerentes dos rankings.

Keywords

Higher education, rankings, research governance, benchmarking, Mode 2 research

Ensino superior, rankings, governança da pesquisa, benchmarking, pesquisa Mode 2

ACKNOWLEDGEMENTS

This project received financial support from CAPES.

To my beautiful wife Janaína, who put up with the stress, shouting and espresso-fuelled late nights with love, patience and understanding. To my parents Claire and Hugo, and sisters Lexie and Tasie, who have supported me so much over the past three years, and are always happy to listen to and reflect on ideas.

I would also like to give special thanks Prof. Jacques Marcovitch, whose drive and support made this project all that it could be, and for going above and beyond what is expected of a tutor, placing a huge amount of faith and confidence in me. I would like to thank also the members of the panel, Prof. Hélio Nogueira, Prof. Elizabeth Balbachevsky whose insightful and diligent comments steered the project into its current form.

To all of the professors and academics who helped form this project, and whose feedback was always very much welcome and important to me. In no particular order; Dr. Carlos Henrique de Brito Cruz, Prof. Luis Nunes, Prof. Simon Schwartzman, Dr. Gordon Finlayson, Prof. Colin Grant, Prof. Rafael Villa, Prof. Enrique Iglesias.

I would like to thank the Instituto de Relações Internacionais IRI-USP for the wonderful and enlightening course of study, and the support offered to me over this time, particularly Prof. Janina Onuki, Prof. Pedro Dallari and Prof. Deisy Ventura.

I would like to give thanks also to Giselle Castro in the IRI office and to Noda Haruko, whose superlative organisational skills and infinite patience were crucial to this project.

Also thanks to Rob, who has listened patiently in the bar for hours to me letting off steam about the project.

Table of Contents

Table of Contents

1. INTRODUCTION.....	6
2. WHAT ARE RANKINGS?.....	10
2.1 Transparency.....	17
2.2 World Class Paradigm.....	18
2.3 Power relations in Rankings.....	20
2.4 Competing Visions of Universities.....	34
3. THE NEW PRODUCTION OF KNOWLEDGE.....	47
3.1 Mode 1 and Mode 2	50
3.2 Triple Helix Industrial Relations.....	59
3.3 Knowledge as a Global Public Good and the Open Access Ecology.....	68
4. RESEARCH AND IMPACT.....	71
4.1 Academic Impact: Journal Impact Factor and Citation Impact.....	73
4.2 Social and Economic Impact.....	83
5. THE SHANGHAI JIAO TONG ARWU RANKING.....	85
5.1 Weightings.....	86
5.2 USP's Performance in ARWU.....	92
5.3 Conclusions from Shanghai Jiao Tong ARWU.....	94
6. TIMES HIGHER EDUCATION AND QUACQUARELLI SYMONDS RANKINGS.....	95
6.1 Times Higher Ranking.....	99
6.2 Quacquarelli Symonds.....	106
6.3 Combined Performance Analysis.....	110
7. THE UNIVERSIDADE DE SÃO PAULO.....	114
8. RECOMMENDATIONS AND FUTURE OPPORTUNITIES.....	120
8.1 Fourth Wave Ranking Systems: Benchmarking.....	120
8.2 New Measurement Metrics.....	121
8.3 Concluding Remarks.....	127
9. BIBLIOGRAPHY.....	130
APPENDIX A: INSTITUTIONAL BENCHMARKING PROJECT.....	140
APPENDIX B: MAPPING CLIMATE CHANGE.....	161

1. INTRODUCTION

This project seeks to answer two main questions on rankings with regards to USP: Firstly, to what extent do rankings represent reliable and objective performance indicators for universities? Secondly, to what extent is the Universidade de São Paulo equipped to respond to the challenge set forth by rankings? In answering these two broad questions, the project concludes that rankings do not present especially clear, attainable or necessarily desirable incentives for modern research institutions. They are however suggestive of a number of crucial features necessary to becoming internationally competitive. The way in which USP is structured and measures information means that although it possesses many of resources needed to reach this level, it is unable to harness them in an effective way. Because of this unsatisfactory double bind, the final sections of the dissertation are dedicated to pointing towards a response to rankings for USP, by looking beyond ranking position towards the creation of benchmarking.

We live in an age of accountability. Public expenditure must be measured in terms of its impact and effective deployment in order to satisfy the requirements of democratic states (RANSON, 2003). One area in which this has had noticeable impact is in higher education, with institutions under pressure to demonstrate effective deployment of resources in a way that has real impact upon society. Because of this, public evaluation exercises have taken a central role in research policy steering over the past decade, nowhere more so than in the rise of rankings and league tables (BALL, 2009; RAUHVARGERS, 2011). In general, this development is to be supported, as it allows higher education institutions to compare and measure performance in a way impossible in previous generations, facilitating a convergence of global institutions and a process of learning by shared experience. However, the nature of the hierarchical ranking has sparked a global 'reputation arms race', (HAZELKORN, 2009) encouraging the adoption of ranking-improving policies rather than institution-improving policies, and attempting to reinforce the hegemony of the established order. Rankings privilege one historical and traditional model of

institution over a diverse and fragmented higher education landscape.

The national and international higher education landscape in which the Universidade de São Paulo finds itself at the beginning of the twenty first century is qualitatively different to that it found itself at its foundation in 1934, its first reform in 1964 or even its second in 1987. The vast increase in both global and domestic enrolment rates means that higher education is no longer the preserve of a small social elite, but a truly public endeavour with a role in creating a knowledge society. The critical public sphere that the university takes a role in fostering has corresponding effects on the institution; the demand for public accountability of the institution presents itself as a new imperative for the institution, whereby the old, school or faculty level 'elitist' forms of evaluation are no longer sufficient to respond to societal demand. In response to this demand, USP began in 1987 to publish an *Anuário Estatístico* (statistical yearbook), giving a quantitative representation of 'the life of the university', allowing the general public to evaluate the output of the university.

On a global level, the institution finds itself inserted into a system that is more globalised, its interlinkages and networks of communication and research are denser than ever before. It has been said that global higher education is on the verge of a 'digital moment' of significance comparable to the effect on publication and information production of the spread of the printing press. This change comes as a result of the vastly increased development of communication technology, from email and the ready availability of conference calling technology, to online publishing and archiving sources and open access collaborative platforms, the internet has allowed the spontaneous and instantaneous collaboration and dissemination of research across national, social and disciplinary boundaries to confront the complex issues facing global and local society (BERNHEIM, 2008; OLSBO, 2013). This is evident from the proliferation of international student mobility, international and transdisciplinary collaboration and the explosion in the quantity of published material available.

This means that the modern university of excellence no longer exists in a vacuum, it can no longer simply reflect state goals or remain inward looking, or take an apolitical, detached view of its activities, like a scientist in a laboratory (HOHENDAHL, 2011; LATOUR, 1987). This nexus of issues,

identified as the transition from an institution operating within an elite system to a mass system means that existing administrative and evaluation systems can no longer cope with the running of the institution (TROW, 1973). It is this backdrop of public discourse and demand for accountability, and the reality of a deterritorialised research landscape, that the implementation of the Anuário and rise of global university rankings can be understood.

As this rise is often conflated with broader processes of globalisation, it has often been understood in terms of economic globalisation and the intrusion of market logic and corporate governance into an area where it is inappropriate. While in part this is true, and many of the current rankings reflect powerful capital interests in a way that is not desirable, globalisation is not one unified phenomenon with unified inputs, outputs, goals and steering, rather it points to a fragmentation and differentiation process that is non-linear in character (APPADURAI, 1996). While economic globalisation may be a worrying factor for many, the exchange of knowledge and capabilities higher education can facilitate on a global level reflects the cosmopolitan, solidarist aspect of humanity at perhaps its most complete level (ARCHIBUGI, 2004; LINKLATER, 1998). The ability to compare and share experiences on an international level is vital to the success or otherwise of higher education systems, allowing a collaborative and cooperative approach to policy formation and learning through shared experience. Rankings are now an inevitable social fact of twenty-first century higher education, universities cannot make them disappear by ignoring them (SALMI; ALTBACH, 2011). This dissertation proposes that they are in fact to be encouraged, but steered away from the structure and tendencies of the current rankings, towards one that more accurately measures impact of research and respects the heterogeneity of institutions and research outcomes.

This paper will first look at rankings in terms of the power relations they represent in order to explain why rankings have become such a central function of global higher education, before presenting a vision of research and the knowledge society based on the new production of knowledge and proliferation of open access to demonstrate that they are at present ill-suited to representing them. It will then look in detail at the individual metrics involved in rankings to demonstrate the ways in which they adequately measure performance, and how

they fail to. This will then be combined with an analysis of the way USP records and measures performance in the Anuário Estatístico, to demonstrate areas of strength in the measuring system, but also omissions that show the gap first of all in the way research is conceived of and the way it is measured, and the way it is measured by the university and the way it is measured by rankings. The final part is a benchmarking project and a series of new possibilities in measurement, using a 'partially overlapping indicators' approach that seeks to 'open out' scholarly measurement at both a macro (institutional) level, and a micro (departmental or issue-based) level (MARTIN; IRVINE, 1980).

2. WHAT ARE RANKINGS?

All types of evaluation or criticism are made in light of a critical standard or ideal. Ranking systems are no different in this respect. When we criticise the government for its healthcare policy, for example, we generally have an idea in mind of what a healthcare system ought to be like; it should be efficient and expedient, it should have up-to-date facilities and drugs, it should be accessible to the whole population and its doctors should be well trained, for example. In the literature this is frequently, and somewhat erroneously labelled the epistemic limitation of rankings. In a hierarchical structure such as a ranking table there is the implication that the top, number one, is the best or closest form of the ideal object in question, while those trailing are less so, because of this it is preferable to label this type of limitation as normative in nature rather than epistemic. It is not so much a problem of how to interpret knowledge and information (which would be an epistemic problem) as much as how the use of a critical standard affects rankings exercises.

This is a characteristic limitation of all hierarchical league tables, in music sales charts, for example, where who has sold most units and therefore is most popular is measured, as opposed to any judgments about artistic merit. Simple quantitative measures are well suited to relatively simple phenomena, but when they are translated for use to very complex phenomena, like the life of a university, in accordance with Campbell's law derived from the use of standardised testing in education systems, they begin to have distortive or deforming effects upon the phenomena they are designed to measure, and subsequently increase the incentive to cheat the metric (CAMPBELL, 1976).

Rankings have the effect of turning vast, multifaceted complex institutions into one dimensional representations. This collapse into one dimensionality is characteristic of all hierarchical rankings, whether they be measures of academic excellence, social inclusion or sustainability. They always necessarily exclude some part or other of a university's function, identity or reality, and always reduce the many qualitative phenomena required to make judgments about performance into simple quantitative data. It is with this in mind that we

should be wary of any of the claims of total representation of quality from a ranking, as there is necessarily something lost in reading universalised statistics.

This one-dimensionality affects all university rankings, and so tends to detract from nations' attempts to formulate alternative rankings when dissatisfied with the outcomes of the mainstream rankings, as Russia did in publishing its own Global University Ranking from RatER which placed Moscow State in fourth, higher than Harvard or Cambridge as a result of its capacity to place graduates in high ranking executive positions (BATY, 2010). For reasons to be detailed below, for critics of the central-periphery structure of rankings or the implicit soft power relations they embody (LO, 2011), rewriting rankings to favour local priorities and redefining global excellence to favour one's own system is not how nations should approach this idea, they should seek to challenge the very notion of hierarchical league rankings, which presuppose a meta-language of research and uniformity of goals (ERNE; 2007).

There are many different variables that could be used to measure what makes a good university; number of trees planted by the administration, wheelchair accessibility, the quality of student life and accommodation or graduate employability scores. All of these clearly reflect something of the Idea of a university, held to some standard of what we feel a university should be. The way in which we choose to rank universities should reflect something of what we think a university ought to be. It is very important to recognise this normative aspect to ranking metrics; we use them as a proxy for excellence, a way of quantitatively measuring how close, or far, an institution is to a set ideal of what a World Class University is.

For the vast array of information that could be recorded about a university, it is a comparatively narrow set of criteria that is used in international rankings, even when compared to domestic rankings (Steiner; 2010). Given the enormous diversity of institution types and structures, priorities in research and learning outcomes and academic traditions there is not really a corresponding diversity in ranking metrics. The weighting is overwhelmingly, although not entirely, given over to research and academic production, rather than teaching or student experience. This suggests that at least when it comes to our global

institutions, there is a degree of consensus present in what makes a university not merely locally dominant, but of global importance (HAZELKORN, 2013).

As Altbach (2004) stated of the World Class University; everybody wants one, but no one is quite sure what one is. It was in the search to define and construct one in China that the first Shanghai Jiao Tong was created (Liu & Cheng, 2013). Because of this a high ranking position is often used as a proxy for World Class status as it is the main aim, at least of Shanghai Jiao Tong, to define those universities at the very top of the global research pyramid.

In the present third wave of composite rankings, the view is complicated by three additional factors; the use of ordinal numbering, whereby statistically insignificant variations in measurement come to have disproportional effects, the difference in position lower down rankings is often so small as to be statistically irrelevant, but can create variations of hundreds of ranking positions, and the composite nature of the rankings, which means that by design, rankings can only favour one type of institutional profile rather than allowing for the possibility that more than one set of institutional goals or models can produce excellence. The scores reported in the three focussed on in this paper are not absolute values, as they are often read, but scores produced in relation the top ranking institution (RAUHVARGERS, 2011). This means that rather than being a progress report as such, it is possible for an institution to improve its performance in some aspect of the ranking, but still suffer a loss of points because the top institution progressed at a faster rate. This gives the impression of a zero-sum competition rather than the idealistic view of the ready and free exchange of information described to which we ought to subscribe (ORDORIKA; LLOYD, 2013).

Composite rankings also combine series of data that are not easily reconciled into one composite score. When governments, or international organisations measure information there is a clear idea of what kind of measurement is taking place. Inputs are measured compositely as the sum of a variety of inputs, efficiency or productivity is measured by pitting the ratio of inputs to outputs, results as the expected outputs, and impacts as an aggregation of the effects of the results. When measuring data we can look at inputs, outputs, intermediate qualities, efficiency measures, and impacts. In

order to have a coherent composite score we should not mix these types of data. The human development index, for example, is composed quite simply from the mean life expectancy at birth, expected number of years of schooling and GNI per capita. This gives a broad impression of the level of human development in a society, and can facilitate conversations on why it is higher in some countries than others where GNI per capita is equal. It does not measure levels of democratic participation, empowerment, liberty, security or inequality, nor are there any plans to expand it to cover these as they are problems of different orders, or contextually dependent. It is simple, transparent and easy to use, but for a representative picture of a population, it must be used in conjunction with a variety of other indicators.

A clean measure should also always avoid double counting problems, wherein the same data are represented multiple times within the score, meaning that they suffer from a multiplier effect. In this, many aspects of rankings are counting the same phenomena repeatedly in adapted ways. Impact, combined with prestige and output into one score is counting the same issue multiple times in subtly different ways.

What we see in composite rankings is a mixture of inputs (industrial resources, faculty to student ratio, and increasingly highly cited researchers), productivity scores (papers per faculty member), outputs (number of papers indexed on a specific source), impacts (highly cited papers) and intermediate qualitative information (reputational surveys). From this cocktail, it becomes difficult to say what a composite university ranking score is actually a measurement of, because it is not purely of inputs, outputs, intermediate factors or impacts. Instead, according to ranking agencies, rankings measure 'quality', or 'excellence', neither of which are coherent or conceptually unified terms for which it is possible to create a single metric.

The time span of measurement for rankings is also problematic, it not clear that an annual cycle is the most appropriate institutional change in a university, which is generally slow moving, and can take years to show the full consequences of decision making or change. For example, if Amazon deforestation were measured hourly, over the course of a week, there could probably be some interesting conclusions to be drawn over behaviour of actors

involved in deforestation; preferring to remove large trees under the cover of higher species, for example, but very little of use that could be extrapolated about overall trends in deforestation. Because of this deforestation is measured over a larger scale that allows the accurate mapping of change. Universities, by their nature and governance structures, tend to be durable but slow moving institutions, with the constant satisfaction and cooperation of a variety of stakeholders and actors needed to effect change. Research programs and departments can take years to begin operating to their potential, and major changes in policy do not typically have instant effects, but generate much longer term results. By publishing annually then, rankings are in danger of capturing statistically insignificant fluctuations and natural variations in output rather than long term changes in performance. Rankings may be able to capture a series of snapshots of university performance, but are unable to describe long term change in a meaningful way. The reason for this cycle is that measurement is not tailored to academic cycles, however long they may be, but to business cycles, which demand annual publication in order to generate revenue for publishers.

This discussion will take two lines of criticism of rankings: first, on the desirability of creating universal rankings with which to measure all universities worldwide. Second, on the possibility of creating one set of performance metrics that can adequately describe impact for all types of academic activity based on bibliometrics, based on an argument formed from Gibbons et al. (1994)'s Mode 1/Mode 2 distinction.

In higher education, any system of evaluation can only grasp a small portion of the whole with any level of accuracy. Of the 17,000 higher education institutions in the world today, about one thousand are evaluated by global rankings, and five hundred are actually listed, and only a hundred have their scores recorded. This means that these rankings are designed to measure a tiny percentage of global academic activity; considered the absolute elite. There are perhaps four institutions in Brazil appropriate for this type of evaluation, of which USP is the principal candidate. As a result, national evaluation systems are geared to measure a much less research-intensive environment. Because USP largely follows these national norms its academic activity is not at present

effectively represented at a high enough level.

Brazil has its own domestic ranking system, the Ranking Universitária de Folha de São Paulo, however it is not of great use to USP as a comparison tool. Any ranking must be geared to a sensitivity that captures the phenomenon measured. Because of this, the vast majority of universities in Brazil would simply fail to register on international rankings at all, as is true for the majority of countries. This means that, as the authors state, the ranking is broadly based on the THE ranking, albeit calibrated to capture much smaller data.

The ranking consists of nine indicators of academic performance (although a detailed breakdown of what these nine are is unavailable online) making up 55% of the weighting, 20% is given over to teaching quality as defined by a reputational survey of academics, 20% of market reputation from a survey of human resources departments and 5% as innovation, defined by the number of patents produced by the university at INPI (the National Institution for Intellectual Property).

The academic survey uses the same format as the QS survey, which simply asks researchers to list the 10 universities with the best teaching standard in the country, while in recognition of the likelihood of human resources directors having less privileged information, they are asked to list just three. This methodology is subject to the same concerns about privileged knowledge that the THE and QS surveys covered in section 3 of this dissertation are affected by; to ask an academic to list 10 institutions would require him or her to have knowledge of a very wide selection of universities indeed, or, rely on previous reputation and hearsay to reflect their experiences. Added to the low mobility of Brazilian academics as a result of the non-transferability of pensions and careers, and the often parochial nature of Brazilian academia, it is extraordinarily doubtful that a majority of respondents have intimate knowledge of the workings of ten different institutions. A detailed critique of the use of reputation scores is provided in the next section of this dissertation.

This, in part explains the extremely poor evaluation of the innovative and progressive curricula presented at UFABC in comparison with its academic output; as a new institution it is not yet broadly well-known enough to register in the academic public conscience, despite attracting international recognition from both other universities and specialists. In this context, it is hardly surprising that USP, as Brazil's first, largest, and most prestigious university consistently and quite comfortably tops this metric. Whether it is a reflection of USP's quality teaching or merely a reflection of general esteem is highly debatable.

Employability of graduates is something quite deliberately avoided by most international rankings for good reason. Employability measurements are very dependent on national context and environmental factors such as the health and globalized state of the local economy and socioeconomic inequality, so on a global level they become less a measure of quality and more a measure of the local environment. The exception to this is the Russian RATEr ranking, which assesses quality in terms of the number of board members of top multinationals a university can count as alumni, although when we talk about readying populations to form a skilled labour force, we are generally not talking only about senior board members, but about the provision of a diverse technical body. On a domestic scale, however, this can be a valuable measure in terms of how well a university prepares graduates for the work environment. Human Resources executives, however, are likely to have even less knowledge or interest of what goes on inside universities, and more interest in halo effects and reputation. This has been part of the centre of criticism of the US News College ranking in the US. To that end, it is again no surprise that USP ranks much higher than any other university in the country.

As number of patents is not normalized per population, or as would be more appropriate, normalized per number of staff working in areas likely to generate patents, USP wins outright, because there are simply more academics producing more work in the university than most others could compete with. Even if it were reported as a 'patents per staff' score, it is not clear that patents generated in house is an accurate measure of a university's innovative output. First of all, it makes large assumptions about national innovation systems (see

section economic and social impact of research), it ignores aspects like spinoff firms, and applied research that leads to patents held outside the university by partners. This measure is somewhat reflective of the lack of this activity within Brazil, as global rankings prefer to use industrial income as a measure of innovation, reflecting to some degree technology transfer and not innovation specifically developed by the university alone.

The composite section of nine academic indicators does not clarify which specifically are considered beyond the number of papers listed on Web of Science, this makes any comment or analysis difficult. However, for number of papers, USP will win through sheer weight of numbers over its national competitors.

While the Folha ranking may provide benefit and insight to lower ranking universities looking to improve, USP is so nationally dominant because of its size, resources (concentration of resources and talent) and reputation that the RUF does not provide an adequate driver for comparison or change. Its position is unlikely to change soon, and so while this may be reassuring to prospective students, it should not be taken to be indicative of much for academics and policymakers.

2.1 Transparency

The main thing that rankings have been instrumental in fostering in higher education is a culture of public transparency in university administration (Sowter, 2013). Following both Habermas (1987) and article 2e of the UNESCO declaration, the modern university, while fully autonomous from the nation state, must be fully accountable to its population. Rankings offer concrete public pressure for institutions to be more transparent in recording their information, and allow for the contextualisation of the information recorded. This means that instead of being presented with a list of numbers in isolation, it is possible to compare how these numbers sit in comparison with universities of similar profile internationally. For an institution like USP this is especially important, as its dominant position on the Brazilian educational and cultural landscape means

that it tops national rankings comfortably each year, and can be led into complacency by national comparison (SCHWARTZMAN, 2005).

2.2 World Class Paradigm

In an attempt to solve Altbach's problem, many commentators have outlined what it really means to be "World Class" as an institution in the twenty-first century. The phrase liberally graces most universities of stature's prospectuses, and administrative and strategic debate is littered with its completely unreflective use. This has led a number of commentators to suggest that the label is little more than a marketing exercise in the neo liberal global 'market' for education; something to entice prospective students and parents without any real substance or significance. Some have even gone so far as to suggest that the entire project, alongside 'internationalisation' is chimerical (HALLIDAY, 1999), a delusional project with little meaning outside of the increasing efforts made by universities to market themselves to attract endowments and more high paying students. This means that in the 'race for reputation' (HAZELKORN, 2013; ORDORIKA, 2010), those who are 'World Class' are those who come top of the rankings, both in bibliometric impact factors and in reputational surveys.

Salmi rightly associates the emergence of rankings as a response to the question of precisely what and where World Class Universities are, allowing a more open and objective conversation about the idea. Salmi identifies three main components to a world-class university, each of which we see, to some extent, reflected in rankings metrics (EDUCATION SOCIETY, 2014; SALMI, 2009):

1. A concentration of top talent among academic staff and students. This means the ability to choose the best and brightest both locally and internationally. Increasingly there is an acceptance that the internationalisation of staff and student bodies is crucial in the definition of a world class, and not merely good, university. This is somewhere that Latin American universities struggle to as a result of their social obligation to open access and a wider percentage of the population. While, as Salmi says, there are concentrations of talent within the university, Latin

American universities tend to favour much more diversity than the World Class paradigm would generally allow, this in turn affects some ranking positions for Brazilian universities. USP however, has a phenomenal concentration of regional talent as a result of its regionally dominant position, putting it in good position to assume World Class status.

2. An abundance of resources is necessary due to the huge cost of building and maintaining a top comprehensive research-intensive university. Brazilian public universities are almost entirely publicly funded, and the recent crises have brought this issue to the fore once again. It is not within the remit of this dissertation to discuss financing, although it is notable that both Salmi's account and the UNESCO Declaration point to the need to revolutionise and rethink financing of higher education, as the costs of maintaining top research and top researchers becomes ever higher. However, the state of São Paulo, through huge FAPESP funding resources is an anomaly in the Latin American context, being both freer of federal dependency than the federal universities of other states and with funding streams closer to OECD averages than to other states in Brazil.
3. World Class universities need favourable governance and a high level of institutional autonomy in order to have the ability to steer policy in a responsive and effective way. This applies to the way in which they are funded and the way in which they are administered. The presence of large-scale bureaucracy seriously inhibits the decision making process, and means that universities cannot formulate and steer research policy effectively. It also specifically states that both academic freedom and institutional independence must be protected at all costs. State intervention in higher education policy has proven to be largely disastrous or counterproductive in higher education. It is in these two aspects that, according to Salmi, USP has limiting factors that prevent it from joining the top table. The presence of an 'assembler' model of governance means that it is excessively and inappropriately democratised and overtly susceptible to party political pressures. This means that executive decision making is constrained by the conservative factions of its teaching staff, and excessively bureaucratic. At the federal level, the presence of evaluation

ill-suited to the evaluation of a leading research intensive university, focussing on quantity rather than quality produces incentives on a national level not replicated either in the university's own strategic goals or in international metrics. These governance hurdles urgently require addressing if USP is to fulfil its potential.

From these three key points, Salmi then draws out the key characteristics of a World Class university;

- International reputation for research and teaching.
- The presence of research superstars and world leading researchers.
- A number of World Class departments.
- Generates innovative ideas and groundbreaking research, recognised by prizes and awards.
- Attracts the best students and produces the best graduates.
- Is heavily internationalised in its recruitment activities and research production.
- Has a sound financial base and a large endowment income with diversified financial base.

Salmi gives a very clear picture of the key determinants of a World Class university, and the challenges for the Universidade de São Paulo, although he ignores the role of geopolitics in his analysis; as the BRICS are slowly discovering, simply investing the money, recruiting the academics and copying the policies of the US is not supplying the results expecting at the speed expected. There are variants on the institutional model depending on political context and relationship of the state to society, but the dominance of the Anglo-American model in definitions in the World Class discussion is pronounced (CHENG; LIU, 2006).

2.3 Power relations in Rankings

The three rankings chosen for this study are not necessarily the most rigorous or most applicable to the Latin American context. They were not chosen on the basis of quality, but rather on the basis that they are the most

renowned and most influential. Rankings are a fundamentally public phenomenon, and it is the ability to wield power and form opinions that gives a ranking power, not necessarily its internal coherence and quality. In order to fully understand the significance of rankings as opposed to other forms of academic and performance evaluation it is vital to understand both their internal design and the social context in which they are embedded (COLYVAS, 2014).

In the past new managerialism, neo-institutionalism and neo-functionalism dominated literature in the field, focussing almost entirely on the process of measurement and performance (PUSSER; MARGINSON, 2013). This has led to the majority of scholarship being relatively uncritical of the rise of rankings as a phenomenon, and tends to disregard the role of politics within higher education studies instead seeking to analyse them in terms of how institutions can improve within rankings rather than whether or not rankings themselves are a healthy development in world higher education. These have also tended to focus on specific metrics as component analyses rather than considering the phenomenon as a whole.

However Marginson and Pusser (2007), Tikly (2004) and Yat Wai Lo (2011) have more recently looked at rankings in a social-critical perspective, as representations of power relations within global higher education. They seek to explain the rise of rankings in terms of hegemony, power, politics, inequality and the proper role of a university within civil society. A modern university, as a result of its central role in determining the social, industrial, and cultural make-up of a nation, is a fundamentally political and politicised institution, it is both an instrument with which to change a society and a site at which changes in society occur. Considering the effect in Brazil of obtaining a degree from one of the prestigious state or federal schools, which can confer great social mobility upon the fortunate individual, but at the cost of the wider population, both narrowing and widening inequality simultaneously (FORGACS, 1935).

In the higher education landscape, rankings legitimate the hierarchy of

the existing order by gearing evaluation metrics towards favouring certain types of neo-liberalised institutions, certain types of 'academic capitalist' publication behaviour (RHOADES; SLAUGHTER, 2010; SLAUGHTER; LESLIE, 1997)

Lukes' tripartite conception of power comprises firstly the intentionality of policymaking, the ability for an actor to exercise agency over the decision-making process (LORENZI, 2005; WHITT; LUKES, 1980). For its ability to account for the way Gramscian cultural hegemony translates into bureaucratic control and agency it is well appointed to discuss power relations in higher education, which do not rely so much on notions of coercion or material notions of hegemony currently favoured in International Relations theory. In Lukes' first pillar we see states having the ability to determine financing of universities at a national level, and administrators determining the distribution of resources at the institutional level. The second pillar of power is the gaining of legitimacy through the equitable distribution of scarce resources (agenda setting). For universities, this means that they will be seen as legitimate if and only if it is seen that these resources are distributed fairly and in the best interest of the actors involved. Finally, if this legitimacy is recognised it is because of the role of ideological factors in place, the special creational myths behind the institution and the Idea of the university. It is this ideological basis that orientates discussions of how the second and first pillars ought to function, and is so embedded as ideology that it is rarely questioned, rather it forms the context of how we ought to criticise and judge the university, rather than an object of criticism itself.

Each individual state and society will formulate its own conjunction of intentionality, legitimacy and ideology over higher education, and therefore prioritise and value its own distinctive characteristics of excellence. Some see universities as the cite of social and critical inquiry, others as the drivers of industry, and others as training centres for the plutocratic elite. The way in which legitimacy is granted in the second pillar depends directly upon the specific view taken of higher education in the third. The issue then, with universities rankings is that they consistently favour one dominant model of neo-liberal capitalist higher education, regardless of context sensitivity. Ranking results become so pervasive within discourse of international higher education

that they form the spine of a consistent narrative about what a university ought to be, giving the impression of a shared game and shared determinants of excellence.

The net effect of this is 'mobilization of bias', wherein over time consensus forms over what excellence is, and therefore the first two pillars become completely subordinate to an external third. In its current form this ideology consistently ignores issues of social justice or openness, also the role of nation-building cherished by Latin American public universities (Ordorika and Pusser, 2007; Schwartzman, 2004). This explains why the Russian RATER ranking measured Moscow State higher than Harvard; it is not a propaganda exercise, but rather the case that Moscow State better fits this tripartite picture of power relations for Russia than Harvard does.

This desire to move away from what is seen as an oppressive hegemony of Northern actors is challenged by the soft power account (LO, 2011), which rejects the oppressive nature of the post-colonial account of rankings. In Asia there has been a concerted effort to subvert Altbach's centre-periphery model by building centres in the periphery based around large industrialised economies. He employs a conception of soft power (NYE, JR., 1990) that rewards the proactive adoption of 'World Class' standards, and so performs a disciplining role (COHEN; FOUCAULT, 1978) on the expansion and globalisation of higher education. On this conception he sees rankings creating a framework in which policy makers willingly and voluntarily 'buy in' to the challenges set forward by rankings; their metrics and definitions of excellence have become internalised and universalised. It sets the agenda for the new deterritorialised research landscape, assuming some of the role previously carried by the nation state in its hard power relationship to the institution, wherein it could steer policy with the threat of the withdrawal of funding. This is characteristic of the Asian attitude towards HE, where there is a perceived need to *catch up* with the West, by competing directly with the centre and closing the gap by adopting the same standards of excellence. In this, we can see that the ideology of the West has been absorbed and adopted by Asian countries relatively unreflectively, with little consideration as to what constitutes excellence beyond that embodied by the Ivy League and the Russell Group.

This is reflected in the creation of the ARWU; the Chinese government wanted to know exactly how far behind they were, and formulate policies to drive a global convergence of higher education (LIU, CHENG, 2005).

This soft power conception can also be read as a neo-imperialist attempt to drive the North's domination over the South by imposing linguistic constraints on research (PHILLIPSON, 1992) and neo-liberal models of higher education favoured by the North upon the South. This is evident in the debates surrounding USP's fall in the 2013 Times Higher rankings, in which prof. Renato Janine Ribeiro of FFLCH (now minister of education) stated;

“Não é da cultura científica brasileira, e latina no geral, adotar o idioma inglês como meio de comunicação entre o professor e o aluno. Não temos como fugir da tradução. Mas isso não deveria ser o ‘divisor de águas’ entre a pesquisa boa e a ruim, como fez a revista inglesa”. (MOURA; AMORIM, 2013)

This effect is somewhat mitigated by the fact that the SciELO network is now incorporated into the Web of Science count, meaning that a lot of Portuguese language work is counted in the THE and ARWU, however it should be noted that there is no integration between the WoS citation reports and the SciELO citation index. This means that for article count the university does not lose out for publishing in Portuguese, but as a citation in SciELO is not counted, it is still at a disadvantage for measures of impact.

There has also been marked criticism of internationalisation from a post-colonialist perspective, citing worries about an isomorphic, stateless and cultureless global higher education, and furthermore one that, because it is produced by the market and shaped by competitive market forces, becomes “its direct and favoured son” (CROSSLEY *; TIKLY, 2004; LEITE, 2010; TIKLY *, 2004).

Many Brazilian commentators are concerned that economic influences are putting artificial pressure on higher education, moving its priorities away from local or regional interests and into those of competition, embodied by international rankings (AROCENA; SUTZ, 2001, 2005). Although USP is a public institution, various factors are beginning, to encroach on what has traditionally been seen as a sovereign area in higher education as a whole. For

example, the QS ranking puts weight on private investment and profitability of research, if this is to be measured by the Anuário Estatístico it could significantly move the research landscape in Brazil.

In a Latin American context there has been much concern on the part of policy makers about the methodological biases inherent in ranking exercises, and the subsequent abandonment of policies better suited to the region in favour of an expensive, zero sum arms race to improve rankings positioning (Dill, 2006; Ordorika and Lloyd, 2011). There are also deep concerns about the volatility in ranking positions, where universities can rise or fall dramatically in a year despite very small changes in data, generating crises in university management where none previously existed (RIGHETTI, 2013) . It has been posited that this volatility is the result of the need to generate headlines every year; as commercial and public enterprises rankings need to generate news and interest from the wider public, and a static or slow moving ranking will not achieve this as much as one where universities soar and tumble each year, giving the impression of a much more dynamic and changing landscape than really exists.

2.3.1 Intentionality

According to Marginson and Pusser's conception, there is a degree of intentionality required in the convergence of rankings metrics in order for them to perform their disciplining role. Some have placed this intentionality with a cabal of elite Anglo-American institutions with a vested interest in maintaining status quo. However, following the Annapolis group's Grinnell statement in 2007, withdrawing from rankings exercises, and increasing critical volume coming from the developed world and established institutions, as well as those on the periphery. There is also talk of a breakaway 'super 20' within rankings, universities who no longer feel that rankings provide a challenge or a driver to improvement. The intentionality on this conception lies not with institutions, but with the ranking agencies themselves.

For all the literature on the utility, function and purpose of rankings, there is surprisingly little published on the role of the actors compiling them.

Ranking agencies can be either public, as with the Shanghai Jiao Tong, Scimago and Webometrics groups, or private enterprises, as with Quacquarelli Symonds and Times Higher. Powerful vested interests for the propagation of the present situation can be found within them. Furthermore, it is worth considering the roles of both the regulatory observatory IREG, and especially the bibliometric information compilers ISI and Scopus, whose information on individual article citations, journal rankings and highly cited researchers form the backbone of all of the major rankings.

2.3.2 Capital Interests of Private Ranking Agencies

Rankings make up a key feature of the new globalised, marketised higher education context. Universities worldwide are now assumed to be playing a common game for prestige, competing for the same elite students and awards with the same institutional goals, and rankings are a reflection of universities' performance in this. Against this backdrop, a multitude of private actors have sprung up in recent years, providing services to prospective students and to institutions looking to attract them. For these actors, principal among them the private consultancy firm Quacquarelli Symonds and the English newspaper Times Higher Education, rankings offer an excellent opportunity for publicity and a way in which to exercise power over higher education. At the annual publication dates of each, the amount of international press generated means that they can maintain high public profiles, giving interviews on ranking performance and generating a market for expensive advertising space on websites, and for even more profitable consultation services to institutions wishing to improve ranking position.

For instance, following last year's Times Higher, head Ben Sowter advised the Universidade de São Paulo through the Folha de São Paulo that its fall was a result of a lack of internationalisation, and that if it wished to regain its position it would have to administer lessons in English. The fact that this constitutes just 10% of the THE's weighting, and it is in the heavily debated reputational metrics that USP appears to have fallen that year means that this observation, while it may well have grounding in truth, is not borne out by a

reading of the published ranking. It also fails to account for the fall to which he was referring, as USP actually increased its internationalised student body that year, although this is not reflected by the ranking.

Quacquarelli Symonds, on the other hand, introduced their 'stars' system, where universities pay to have extra institutional information included. This is represented in the form of stars alongside the normal ranking, giving the impression of official endorsement rather than the commercial service it actually is. This was the one published recommendation from the IREG audit, recommending the separation of commercial and ranking services.

2.3.3 State Actors

If every nation's higher education policy is intrinsically linked to its foundational myths, stated goals and Idea of a university, it then follows that rankings compiled by these state actors will, to some extent, reflect those same interests (MARGINSON, 2010).

The Chinese government views its universities as the driving forces behind state industry and innovation, and as a result privileges technological capability above other areas, especially social sciences and humanities, as the formation of a critical citizenship is not a national priority for the Chinese government. Because of the remit of the original project, the Shanghai Jiao Tong today shows a very heavy bias towards STEM subjects (science, technology, engineering and mathematics) over all others. The role of teaching is completely neglected, in part due to the near impossibility of identifying sound methodology to cross compare teaching environments, in part due to the fact that China's priorities are in building competitive scientific capabilities for the state and not in providing a consumer guide for prospective students, who will find little of practical interest in describing student experience.

It is often declared in the literature that rankings consistently favour the neo-liberal market model of university over public institutions like those found in Latin America, however we come to a puzzling anomaly in the Shanghai Jiao

Tong ranking, which as the product of a Chinese governmental research program into excellence in higher education, reflects the interests of the Chinese government, which exercises both a highly rigid control over its higher education and over its economy. Some commentators have been tempted to say that it is the product of Anglo-American soft power over agenda setting that has led it to consistently privilege these models (LO;2011). The truth, however, is more complicated as the Shanghai ranking is formulated to the interests of the Chinese government, which is markedly not the same as the Anglo-American model.

The Chinese have consistently neglected the value of critical, liberal education for both ideological and historical reasons, preferring focus on exact sciences as a means of extending state interests. To this end, the governance of the Chinese key universities is very closely controlled by the Chinese Communist Party (CCP), and the ARWU was developed as a tool to assist the use of universities as drivers of national development. In this, its interests in fact align closely with the Anglo-American model with an emphasis on exact and applied sciences for their economic and innovative potential, but not with the European and Latin American focus on a rounded liberal education of its citizenship.

In this way we see the reflection of two very different types of state reflecting the same kinds of interest in opposition to a third, very different. Because the United States is a liberal market economy, its universities are geared towards a demand for revolutionary, short term innovation, with lower regard for solidarist conceptions and liberal education, while European economies tend towards more coordinated market economies, favouring liberal and comprehensive education and slower, longer term innovation cycles rather than huge paradigm ruptures (THELEN; 2012). The Chinese, while not a liberal market economy, have directly tied their industry and manufacturing to innovation rather than to general competency, and so the demand for specialized skills in China is low, as their economy becomes more mechanized, as opposed to in Europe, where it is very high. Because of this, rankings tend to represent rare and exceptional research rather than general competency, and privilege US and Asian institutions over European and Latin American ones.

The Webometrics research group, by contrast, focuses on the web presence and cultural extension of research, reflecting the Ibero-American preoccupation with knowledge as a public good, one of the central concerns of the Havana Conference of 1996. As a result of the way knowledge is viewed by the two states, one as a key driver of state power, the other as a public good to be shared in a process of common learning, the two rankings are structured completely differently. While ARWU measures impact in terms of the academic influence it has had, Webometrics measures it in terms of its extension and dissemination, attempting to give an account of social impact (THELWALL; VAUGHAN; BJORNEBORN, 2005; THELWALL; VAUGHAN; BJÖRNEBORN, 2005).

2.3.4 IREG Observatory

In recognition of the fact that rankings exercises have an influence far beyond being simple curios, and impact on all levels of higher education governance, a group of rankings experts and higher education professionals founded an international observatory of rankings exercises in conjunction with UNESCO-CEPES to ensure that the same standards of transparency expected of the universities they measure can be expected of the rankings agencies themselves. Coming from the preliminary meetings, the observatory published a charter for ranking organisations to follow, the Berlin Principles (UNESCO-CEPES, 2006). Its aim is to attempt to regulate ranking organisations' behaviour and counteract some of the criticisms levelled at the global rankings by the academic community.

The principle recognises the need for rankings to become more responsible and aware of their impact upon the higher education landscape. The focus of the first section of the principles is on rankers being clearer about the purpose of their measurement, not to make their rankings look objective but to be clear about why they are measuring, with special attention paid to representing the diversity of institutional types, and especially national priorities and contexts. The second is concerned with transparency of methodology, on

the basis that any ranking is a useful tool only to the extent that its workings are clear, intelligible and marking to a generally acceptable critical standard. The principles put forward preference to measure outcomes wherever possible as a mark of a university's actual performance rather than inputs, which may only serve as a measure of the condition of a university or to the capriciousness of its benefactors or selectivity of its entry procedures. In practise, this is rarely strictly adhered to by commercial rankings for reasons set out in the next section. The third applies to the application of ethical standards to data collection, being clear and responsible in the way data are collected, and the fourth to the honest presentation of results without distortion or non-disclosure.

IREG's stated function as an observatory is to ensure rankers maintain high standards by offering a currently voluntary auditing process. At the time of writing, the only global ranking audited was the Quacquarelli Symonds, which proudly displays its IREG accreditation on its front page. However, what the auditing process consists of is yet to be made public, as have the results of the auditing process itself. All that is available from IREG's website is a summary of the final decision, stating that it finds the QS methodologically sound and compliant with the Berlin Principles.

As mentioned before, IREG's executive board is composed of experts in higher education evaluation, including Robert Morse. Morse is a journalist and the author of the influential, but highly controversial US News and World opinion rankings that prompted the Annapolis group withdrawal (BENNET et al. 2007). Following the break between QS and THE, Quacquarelli Symonds was acquired by the same publishers as US News, of which Morse is also editor-in-chief and now also sits on QS's advisory board. This means that one of the executive board members responsible for deciding the formatting of the audit is also has a hand in formulating the ranking, and is making a profit from its publication. This potential conflict of interest in granting IREG accreditation has not been widely reported, and it is unclear as to the level of involvement of Morse had in the auditing process, but developments such as these are unlikely to allay the concerns of members of the academic community vis-à-vis the power structures and biases propagated by rankings.

The author has attempted to contact IREG for details of the auditing

process, the make-up of the auditing team and the detailed results of the QS audit, but at the time of writing no response has been received.

Until the auditing process is made more transparent, and the findings of the audits published in full the suspicion will remain that far from becoming a neutral observer in the rankings process, the observatory will in fact become an uncritical legitimating influence on the rankings process as a result of the composition of the executive board.

2.3.5 Citations Compilers

ISI, which collates data for Shanghai Jiao Tong, Times Higher and Scimago from the Sciences Citations Index- Extended (SCI-E) and Social Sciences Citation Index (SSCI) is the intellectual property division of Thomson Reuters, which through its publishing and intelligence units reported in 2013 turnover of US\$923 million. Scopus is the intelligence arm of the publishing giant Reed Elsevier, which in 2013 reported profit of £1.794 billion, from turnover of £6 billion. Both companies feature centrally and profit heavily from the system of 'academic capitalism' that treats knowledge as a commodity produced by the university-as-factory (TALBURT, 2005). This means that both have a strong vested interest in mobilising pressure against open access initiatives by over-valuing paid journals, as open access sources, like SciELO or the university's repositories work against the financial interests of the agencies collating ranking scores (WELLEN, 2004).

ISI and Reed Elsevier also have strong incentives to favour their own journals in journal impact factors, although this point has never been proven either way (MARGINSON, PUSSER;2013), the lack of clarity over methodology in journal impact factors, showing them to be elaborate bargaining processes rather than externally verifiable sources (BUTTON et al.; 2013), means that the potential to use bibliometric citation information to encourage academics to publish in paid journals rather than open access ones is clear. Because of the size of the data sets used by these agencies, there are inevitable discrepancies between their data, however quantitative research into the two has thus far proved indeterminate, some demonstrating higher citation rates from ISI, others

from Scopus. Most agree that these two are the most comprehensive and reliable compared to other resources like Google Scholar and for this reason they tend to dominate the rankings landscape.

2.4. Conflict with the UNESCO Declaration on Higher Education for the Twenty First Century

UNESCO's 1998 *World Declaration on Higher Education for The Twenty First Century: Vision and Action* lays out general principles for the development of higher education suited to the twenty first century. It does this on the basis of the unprecedented increase in demand and diversification of the sector visible in the movement from elite to mass systems of higher education, and in recognition of the fact that in accordance with the Universal Declaration on Human Rights Art. 26 paragraph 1 that 'Everyone has the right to education'. It emphasises the need to open and fair access to higher education (art. 3), something which all of the major rankings have pointedly disregarded in favour of promoting very selective admissions policies (SALMI, 2009; ORDORIKA & LLOYD, 2011).

It also promotes the advancement and dissemination of research (art. 5), a long-term commitment to research relevant to the societal context in which it is located (art. 6), again something disregarded by both composite and bibliometric rankings (ibid.). Perhaps most importantly, to qualitative evaluation of excellence in universities (art. 11). This section specifically states that 'due attention should be paid to specific institutional, national and regional contexts in order to take into account diversity and to avoid uniformity' (art. 11a). This clause would appear to specifically discount the value of a global ranking if we agree with Marginson and Pusser's argument that rankings continually promote and encourage one model of university and one type of excellence regardless of context, which also finds problems in article 8a, which promotes a diversity of models to satisfy the greatly diversified demand for higher education. This observation has been reinforced by Steiner's (2010) meta-analysis of the main ranking methodologies, finding that they are, by and large, uniform processes

with surprisingly little variation between them. Key determinants are measured in broadly the same way by all three of the main rankings, and it is because of this it tends to be the same cabal of winners at the top.

Rankings exercises do not fit especially well with the aims set out by UNESCO for the future of higher education, and yet, to co-opt the language of global governance, they have become extremely powerful informal agenda setters in the higher education landscape, with some 90% of executive decision makers confessing to making policy decisions in order to obtain better ranking position (GIACALONE, 2009; HAZELKORN, 2009). Some commentators have chosen to explain this in the psychological terms of the urge to compete and compare (SOWTER, 2013), while others have chosen to equate the exercise with the increasing incursion of market logic in the higher education sector (MARGINSON, VAN DER WENDE; 2008). These are all contributing factors to some extent, but this analysis would like to take into account the structural and historical factors that explain why ranking exercises have gained such importance, and also where these impulses may be used alongside the preservation of institutional, national and multilateral goals.

At present, I contend that global universities rankings are not fit for purpose either in relation to the normative vision set out by the UNESCO world conference on higher education or present understanding of knowledge formation or the research process. By capturing a limited range of quantitative factors and presenting them as holistic indicators of quality they steer policymakers into making decisions not necessarily in the interest of the institution.

Art.2e states that universities must 'enjoy full academic autonomy and freedom, conceived of as a set of rights and duties, while being fully responsible and accountable to society.' It is within this concept that the value of rankings may show itself; in order to be fully accountable to a society, a university must be able to show what it has produced in return for its investments, and is

contributing to society. These two points, although separate are fundamentally interlinked (GIBBONS et al.; 1994). In other words universities must be able to show the impacts produced by research stemming from the input of public funding. The ability to do this across national contexts allows decision makers and populations compare and contrast the impact of research much more easily allows local populations assess the quality of university in a broader and better informed way than would be possible without them.

The intention of this project is not to claim that rankings do not correspond to one particularized vision of the perfect university, rather to demonstrate that as a university is such a multifaceted organization that any attempt to present one version of excellence, quality or performance is bound to fail, or at least fail to represent a variety of important aspects of the university. None of the conceptions below is privileged above the others, or meant to be preferred to others, and the analysis of the effect of rankings on the balance of them, specifically in Brazil, is to show that rankings represent the artificial introduction of environmental factors not necessarily aligned with Brazil's.

2.4 Competing Visions of Universities

It is notable that at no point in USP's founding decree or any subsequent documentation, is there mention of the university's exact purpose or aim. One might surmise that this is indicative of a lack of direction or leadership. In reality, this is not an omission, but rather a reflection of the fact that there is no single purpose of a university, but a variety depending on the perspectives of different stakeholders. What we find instead is a series of guiding principles to which the university ought to be upheld.

Olsen (2005) demonstrates this heterogeneity of institutions not just between institutions, but also inside them. First of all, he distinguishes between two specific modes of the university, the instrumental and institutional before outlining four different ideal types, each of which find their expression to some greater or lesser extent in all universities, determined by a nexus of internal, environmental and historical factors. It is through this analytic lens I would like

to focus the analysis of USP and rankings, in order answer the question of what, exactly, is gained and lost through the use of rankings to measure performance.

2.4.1 The University as an Institution

The university as institution refers to the internal rules and traditions that constitute the university's identity. History, foundational myths and principles, and the identity of actors compose the university as an institution. The identity of professors, undergraduates and doctoral candidates both encourage and constrain certain behaviours that correspond to their respective definitions. In other words, it is the Idea of the university referred to by Habermas in the great reform debates in Europe during the 1970s (HOHENDAHL, 2002). Olsen contributes two visions of the university as an institution; the university as a community of scholars and the university as a representative democracy.

No institution is a pure representation of one of these ideal types at the expense of all others, and they are by no means the only possible ideal types of a university. This analysis is intended to show that first of all universities are subject to competing visions of higher education's purpose, and second of all the balance of these competing visions is subject to a variety of factors.

2.4.2 The University as a Community of Scholars

The identity of a university as a community of scholars finds its roots at the very heart of the establishment of the modern university with the *Habitus Authenticus* marking the foundation of the University of Bologna, which guaranteed scholars immunity from local affairs, including the right to be tried separately in ecclesiastical courts and not local ones, and immunity from local political affairs (KIBRE; 1954). Through the classic Humboldt model of university, which prescribes the absolute union of research and teaching, and total academic autonomy, this has morphed in the modern era as a conception of the self-regulating Republic of Science (POLANYI; 1962) which manages its own affairs internally, and generates its own definitions of success and internal

evaluation, regardless of external standards. In this way, we can see that much of the debate over Mode 1, homogenous basic research falls into this category, as academic disciplines are largely defined and determined by the academic community, and not by societal or economic necessity.

It is through this identity that individual academic autonomy is derived, one of the sacred principles of higher education that ensures that academics have the right to complete freedom of study untethered to the interests of external actors, such as political parties. This is undoubtedly a necessary component of research, the ability to be freer from the demand for instantaneous or tangible (economic) results for the pursuit of deeper knowledge across a wider variety of knowledge necessary to culture and development. If this aspect is not present at all, the university becomes indistinguishable from a corporate research and development department, and loses this specific advantage. It is also especially important for the humanities and many social sciences, which do not have instantly applicable results but are nonetheless absolutely key to cultural and political life. This conception has come under significant attack since the second half of the twentieth century as the use of the university as an instrument becomes ever more important and influential as the cost, scope and importance of higher education has grown in modern societies.

To a certain extent, this pure community of scholars takes the form of a self-perpetuating myth from the 19th century, scholars eager to ring-fence newly formed and compartmentalized disciplines from over eager benefactors and interfering governments. The Humboldtian principle joining research to teaching, and fundamentally separating the university and state can be viewed as a product of Von Bismarck's enlightened despotism model of government, and therefore perhaps should be viewed in terms of historical context and not taken axiomatically.

The self-governing republic is often taken as a historical image of a past era of higher education.. It is the basal image of Slaughter's academic capitalism concept, and many others writing, largely from the humanities. As attractive as this image is, the truth is that this academic freedom has always

been constrained by one simple factor; the need for resources. In order to elevate higher education funding to above that of philanthropic funding for the arts, universities have always been and will always be dependent on the production of results, whether they are in teaching or in research. Whether for the development of agricultural and navigational tools in the 17th 18th and 19th century, contribution to the military industrial complex for much of the 20th, or contribution to industry and human development in the 21st, research has never been quite as disinterested or depoliticized as many scholars would like to suggest (LEYESDORFF, 2001). The tendency has grown in recent years because of the increase in public attention and demand, quantity of funding and industrial entry, but to suggest that this is somehow a new invention is a misrepresentation of the past. As research and research programs become ever more complex and expensive to fund, this reliance on producing results for stakeholders will only grow.

The QS and THE rankings attempt to represent this republic of science by way of their reputational surveys, an attempt to show the internal regard with which an institution is held. However, as shown in the analysis in this dissertation, reputational surveys tend to reinforce the negative aspects of the self-governing meritocratic society; the inherent conservatism, appeal to authority and historical factors over actual performance and competences.

2.4.3 The University as a Representative Democracy

In USP, and in Latin American universities as a whole, the ideal that a university should be an open, discursive place where its governance is fully representative of its actors gains enormous power, especially amongst faculty members and student bodies.

The debate over universities as democracies gained enormous force in European debates during the radicalism and reform periods of the 1960s and 1970s, during the time in which the chairs system was abolished and there was a concerted effort to tear up much of what was considered a broadly discredited system that was outdated, ultra conservative and perhaps most importantly, had

proved itself unable to resist and in some cases collaborate with despotism and fascism. The idea that the university must become a discursive and inclusive space, and represent egalitarian and representative democracies to be replicated in the broader society became key demands for younger faculty and student bodies. However, from the 1990s onwards, as universities have tended increasingly towards more corporate management models and the balance swaying towards the university as an instrument of service (see below), the resistance has been, on the whole, understated when compared to that experienced by USP in recent years.

It is possible that the status of a national democracy may act as an environmental factor in determining the strength of support for representative democracy in universities. While European nations were emerging from dictatorships and installing incipient constitutional democracies, which by their nature are fragile and subject to low public confidence, support democratic universities was at its peak. As national democratic systems consolidated and stabilized, and public confidence in them grew, the debate gradually took on a less important role in higher education debates. In the US and UK, the two longest established and arguably most stable democracies, this ideal type has never been a central issue in debates with some form of stakeholder representation preferred usually more in the interest of managerial expediency than ideological bent.

Brazil, meanwhile, is only twenty years in to full constitutional democracy following military dictatorship, its democratic institutions are still nascent and public confidence in democracy is still fragile. It was only in the 1980s amid a scenario of stagnation and dissatisfaction that the university began to undertake a democratization project seen as necessary at the time to revitalize the institution (GOLDEMBERG; 1996) Similar progressions can be seen in the evolution of UNAM in 1946 and then in 1967 (ORDORIKA, PUSSER; 2002). This hypothesis could be the subject for further research. However, it suggests that university governance systems are not formed accidentally or independently of national context. It also suggests that, contrary to the implicit suggestion of rankings, that wholesale importation of Anglo-American

governance structures to Brazilian higher education would produce world class Anglo-American style institutions.

Rankings do not measure the democratic or representative nature of universities in any of their metrics, as these are not treasured aspects within the Anglo-American orthodoxy, nor are they reflected in the goals of the Chinese government (ARWU). Indeed, they do not take governance structures into account at any point. If as established by March and Olsen, one of the main things to be gained from international measurement is to better understand the likely outcomes that different institutional structures and arrangements have, this a way in which benchmarking systems can begin to further understanding beyond the superficial, reductive way in which rankings represent performance.

If it were possible to understand the differences in outcome, for instance from an appointed board of governors as opposed to an internally elected executive, or how different levels of democratic participation affect academic performance, benchmarking could have the power to draw much more profound conclusions about university governance. The University of Cambridge, for example, incorporates an electoral process into its decision making with the Regent House comprising 3,800 resident senior staff and officers. They elect the chancellor (a symbolic role), and also the majority of the university council, composed of sixteen senior academics to oversee the majority of administrative issues within the university. This council then appoints a special council to select candidates (usually three) for the vice-chancellorship based on a seven year fixed term. The candidates are interviewed by the council of sixteen, and the appointment is then put to the Regent House for approval. Although this is democratic in theory, in practise, the Regent House rubber stamps the selection of the Council.

This system is much more open and representative than most US models, and more in line with USP's, however by ringfencing financial issues and various other aspects of the university, it is perhaps better protected from mismanagement than our institution. (UNIVERSITY OF CAMBRIDGE; 2009)

The University of Oxford runs a similar, theoretically electoral system for choosing its executive; a list of candidates is selected by a special committee and then chosen by the University Council, composed of sixteen senior representatives who compose the main executive function of the university. This candidate is then approved subject to the approval of the congregation, composed of 4,500 senior and retired academics and administrators (UNIVERSITY OF OXFORD, 2002; 2015) It should be noted that there is no student, junior or union representation at any level of executive selection of Oxford's process, so while we could term these broadly democratic systems, they could not be seen as truly representative democracies. The university is currently undergoing a selection process to replace its vice chancellor, electing a woman for the first time, Prof. Louise Richardson. Prof. Richardson is an Irish academic of the first generation to be university educated, and so possibly an example of how the council is managing to change the traditional elitism of senior academics in elite institutions (ORDORIKA, PUSSER; 2002) paradoxically by being less democratic.

By contrast, Imperial College London follows a much less democratic model; there is no congregation. Of the council, one representative out of 16-19 is chosen by election from among the staff, while the rest are appointed by the council itself. The general population has no input into selecting the president, and no formal recourse to remove the council from position. This model is more typical of US and private institutions, and demonstrates that there is significant variation in management models, even within the same academic culture, and at institutions of the same type (IMPERIAL COLLEGE, 2007).

It should be noted, furthermore, that Vice Chancellors and Presidents from these institutions are usually chosen from outside the university, in comparison to Latin American universities, which usually stipulate within their statutes that the rector must be chosen from within the body of docents. There are a variety of side effects to this tendency; often unpopular vice-chancellors are accused of not understanding the institution, or of not having its best interests at heart. There has also been a rise of the professional senior administrator, who moves from large institution to large institution (the current VC of Oxford is a former provost of Yale), giving rise to a 'transfer league' of

high flyers with a corresponding explosion in senior administrator salaries compared to that of academics. On the other side of the debate, universities can have access to the most competent candidates available regardless of nationality or educational background. One of the major problems with the USP model is the need to constantly mobilise political interests and build internal coalitions of support, which then in turn expect favours. Candidates from outside the university are likely to have forged fewer political alliances than internal candidates, and are therefore less dependent on factional support than the reitor is in USP.

2.4.4 The University as an Instrument

Looking at the university as an instrument is to treat it as a means to an end, a tool of instrumental rationality. This means that its outcomes will generally be tangible and measurable. It also means that it is an instrument *for* someone and *for* something. To this end, Olsen distinguishes two different modes of the university as an instrument; as a tool to service markets and as tool of national interest. It is for the university as an instrument that rankings find their most powerful expression.

2.4.5 The University as a tool for the Advancement of National Interests

As higher education has the ability to determine almost every aspect of a nation, from military power (through research), social equality, economic strength through both innovation and equipping an educated working population to a variety of other benefits, it has become central to governmental planning. As such, it is increasingly used as a tool to advance national goals, with increasing governmental steering through allocation and removal of funding to constrain policy choice and control behavior. Rather a the free and meritocratic community of scholars, the university is an extension of government which focuses on problem solving and the fulfillment of the political agendas of national governments.

National governments can exert steering and pressure on universities to produce results in a variety of ways, from direct control of policy and selection of executive, as in the Chinese system, to the imposition of quality assessment exercises in most of the rest of the world, tying attainment of goals directly to institutional funding in a hard power counterpart to the soft power exerted by rankings exercises (LO; 2005).

2.4.6 The University as tool for Competitive Markets

The increasing entry of corporate principles and focus on industrial development is currently one of the most hotly debated topics in higher education today. In Slaughter's (2005) book *Academic Capitalism* she argues that as public funding for research has dwindled in the years following glasnost, universities must increasingly seek funding from industrial partners. Consequently, research is less curiosity driven, and more market driven, to the detriment of areas of low economic impact like humanities, and on both pure research and research with societal rather than economic impact.

As universities' potential to be such key determinants in economic development (OECD; WORLD BANK;), for many areas of research, especially in more applied areas, industrial crossover is not only a reality but a necessity for modern universities. The frequent presentation among the academic community (usually those outside of these applied areas) of private money in public universities as a de facto negative influence on higher education, and a threat to long cherished independence is as simplistic and reductive as those who favour complete marketization of higher education. Industrial insertion forms an important dimension of a modern university's functioning, but only in certain areas.

Slaughter points to the rise of tuition fees over the past decade, turning the role of the student from a scholar expected to contribute to a public good and the growth of learning into that of a consumer who parts with money in return for a professional advantage in the labour market. It is in this that she also points to the turning of the academic prestige of the university as a

community of scholars into a commodity, whose value is defined in terms of marketability. Nowhere is this clearer than QS and THE's use of reputational surveys, which transform academic prestige into a quantifiable score to be compared alongside bibliometric production. It is used not as a mere reflection of how the university is perceived on the outside, but as a constitutive part of the university's 'brand identity'.

The prominence of the competitive market paradigm has thrown the concept of both competition and competitiveness to the fore of evaluation of higher education. While it would be folly to suggest that capitalist interests invented competition amongst institutions and academics; academia has always been adversarial and competitive by nature; between competing schools of thought, orthodoxies, academics and institutions, and during the Cold War between competing political ideologies. What is different in market competition is that competition is now quantifiable, and therefore it can be ranked. Rankings consolidate and reify this competition, performing a disciplining role (MARGINSON, PUSSER; 2001) creating a conception not of competing models, but competing units within a shared game.

2.4.7 Competing Visions of Autonomy

This competing visions model is also demonstrative in the effect it has had on debates over institutional autonomy, as the conception of autonomy and its justification is markedly different for each. For a community of scholars, autonomy is axiomatically guaranteed through the principles of free investigation and inquiry, this autonomy is given to individual scholars to pursue an area of research without limitation from government or institution. This means that any attempt at quantitative measurement, as in rankings, becomes an absolute constriction on this autonomy. For this, assessment criteria of success will be internal; of scientific quality and generally carried out via processes of peer review and evaluation.

For a representative democracy, autonomy is derived from the demand for workplace representation and a fair ability to decide how resources are allocated. This is an autarkic conception of autonomy; it is freedom from outside interference of any kind in the running of a university (OLSEN, 2005). Therefore measurement may offer a threat to certain parts of the university, and a benefit to others, depending on the discipline and the strength of the department. Differentiation and stratification exercises often run contrary to the interests of democratic governance, as there are often winners and losers not necessarily tied to popular opinion within the faculty. Democratic allocation of resources is unlikely to coincide with allocation of resources based on need or scientific evaluation.

For the university as a tool for national interests or competitive markets, autonomy is conceived of as in the hands of the executive, which has the ability to implant policies and carry out predetermined objectives, whether political or financial, using the university. Autonomy is justified here not as axiomatically given, but either for efficiency and expedience, as it is widely recognized that excessive governmental intervention does not produce desirable results (SALMI, 2002), or as a method of increasing stakeholder responsiveness by managing and steering research programs into having significant results for industrial or social partners. It is here that performance analysis through rankings, or more properly benchmarking, can have a real benefit in allowing decision making to be better informed, and the consequences of decisions be better understood.

Federal funding for public higher education in Brazil has not, as of yet, experienced the same decline as seen in the UK or USA, and as a result the push towards industrial entry of research has been much despite gradually declining resources more focused around a few key areas. Similarly, as USP does not rely on tuition fees to supplement income, it does not need to insert itself into the global marketplace for students in the same way as British institutions do out of a growing economic necessity.

Most crucially perhaps, is that the continuing public funding of higher education in Brazil means that Brazilian research is still geared towards public

goods rather than economic results. This is something markedly missing from all of the global universities rankings, which focus principally upon academic impact, and secondly on economic impact.

On the other hand, the absolute dependency Brazilian HE has on state money means it is inextricably tied to state goals. For USP this is not necessarily always a positive thing. According to an interview with prof. Luiz Nunes, president of CERT Comissão Especial de Regimes de Trabalho) Neither international nor RUF rankings exert as big an influence on policy within the university as the CAPES evaluation and dependency on CNPq scholarships for post-graduate support, even despite the more qualitative internal assessment processes and enormous investment from FAPESP. The triennial quality assurance exercise seeks to evaluate the scientific production of Brazilian HE department by department by creating strictly bureaucratic quantitative metrics to measure courses, which in turn has a massive effect on funding. Because of this, the quest for the conceito 7 grade (world class), is one of the dominant goals in the Brazilian higher education landscape at the expense of other, potentially more important goals.

While this has been important to Brazilian higher education as whole in consolidating a research base and ensuring minimum standards of production, as a quality assurance program should, the uniformity and rigidity of evaluation is having deleterious effects on elite higher education's ability to fulfill its social, academic and economic role (CHAIMOVICH; 2000), instead focusing on the satisfaction of unhelpful indicators to ensure survival and success.

A large proportion of the CAPES evaluation is given over to academic production, not impact or any other outcome. This production is defined rigidly, in terms of articles and books published, until recently with scant regard as to where these papers are published, or the quality of them. This means that 'publish or perish' for Brazilian academics is not related to the journal impact factor arms race, as it is for much of the world, but to publish as much as possible.

This has several side effects; the first is in the vast number of articles produced by USP, as researchers are under direct pressure to publish more, not better. As impact is not a motivating factor, in many cases quality will suffer as a result. More damagingly, as the CAPES system is geared towards the generation of short term results, research becomes the constant harvest of low hanging fruit. Long-term, high risk research is actively discouraged by this kind of measurement, and it is exactly this kind of research which has the potential for extremely high impact, revolutionary results. Similarly, as the evaluation is broken down by discipline, it actively discourages the formation of trans and multidisciplinary research, as the researcher is forced into traditional boundaries, or into following multiple Qualis (guidelines) simultaneously (AXT; 2004)

The Qualis have been altered in recent years, with articles stratified according to their journal of publication, according a higher score to articles published in journals with higher impact factors. Problems with using journal impact factors as proxies for impact itself notwithstanding, this can be seen as an improvement in the way in which CAPES conceives of scientific production. In this way, there is more incentive to publish better work, rather than large quantities of low impact research, although the incentives may still be somewhat loaded in the interest of publishing two or three 'low hanging fruit' rather than one, outstanding paper. It remains to be seen whether or not this has changed publishing behaviour within Brazilian academia, as such a sea change may take several years to take effect.

There is still the danger, expanded upon elsewhere in this manuscript, by privileging certain publishing types over others, that this creates a rush towards A1 ranked journals. If academics will submit papers to certain journals regardless of whether they are the most appropriate home for their work, this in turn creates long delays in the publishing process and higher rejection rates.

As the Qualis are discipline-specific, for some disciplines like environmental science and engineering they incorporate assessments of social insertion and industrial insertion. However, at present although these efforts are

obviously extremely promising and contribute to moving away from a scenario where only academic impact, or JIF is valued, they lack quantitative or external metrics for measurement. This means that social and economic impact is assessed in a purely subjective and qualitative way, pointing to a lack of formalised metrics and therefore detracting a little from its potential influence. Mapping or recording social and industrial influence would greatly assist with assessment, and provide clearer incentives and goals for departments.

In recent years, in many developed countries, internal evaluation has moved towards a more qualitative approach in order to allow a greater degree of flexibility to follow exterior goals, however this has not yet happened in Brazil (with the exception of FAPESP and a few other institutions in CONFAP), and while these are not without their critics, they do theoretically permit a greater degree of liberty. In our interview, prof. Nunes suggested that perhaps the easiest and most effective way of changing this in Brazil would be with the creation of a committee to assess at least special and marginal cases to allow higher risk, longer term projects that do not necessarily produce several articles per year for their duration, but at the project's conclusion. This would avoid cases like the members of the genome project losing their CNPq scholarships mid-project because the complexity of the project meant that they were not ready to publish during the first year.

As long as USP remains so far in thrall to CAPES and CNPq, it is unlikely that it will make significant progress within the international rankings analysed here. Isomorphic quality assessment exercises like CAPES that focus on scientific production above all else produce strong incentives counter to those propagated by rankings. One of the university's biggest competitive advantages over private research and development is that it is not tied to any rigid time constraint or demand for instant results, allowing for deeper analysis of an issue or problem, this advantage is dramatically reduced by the imposition of publication quotas with no consideration of the content being produced.

3. THE NEW PRODUCTION OF KNOWLEDGE

The nature of the relationship between the research community and

society has been significantly changed in modern times by a combination of convergent factors. These combine to produce a qualitative shift in the nature of research and evaluation that can be seen as the 'breaking down the walls of the institution'. This has placed serious pressure on the Latin American system in transition, where old paradigms of orientation and measurement increasingly find themselves lacking or struggling to cope with their new functions. Research under the old elite system of higher education focused on specific sites of research and researchers producing work within set disciplines, from which work would then be distributed to a narrow specialised public, little attention is paid to relevance and impact of research, but heavy emphasis is placed upon volume.

Much of the material on the new production of knowledge (NPK) is highly contested. It has been argued that much of it is not new at all, for many fields it is clearly inappropriate or irrelevant, and like much of the discussion of globalisation, many of its diverse effects often appear contradictory or paradoxical (SCHWARTZMAN, 1993). Furthermore, discussion of NPK often confuses the merely descriptive, as in the majority of Gibbons et al.'s 1994 account, with the explicitly normative, saying not only that research is changing, but also that it ought to change more. My intention here is not to give a full account of this debate, but offer one normative reading of it in order to give an account of how Latin American public universities can fulfill their societal responsibilities while maintaining their autonomy, in line with the UNESCO Declaration (1998).

First of all, is the vast expansion in higher education enrollment and demand, from a restricted set of professional elites to a much wider socioeconomic pool, which has led to the formation of a much wider discursive public sphere for higher education (GAUCHAT, 2012; HABERMAS, 1987; TROW, 1973). This in part explains the rising demand for evaluation that rankings satisfy by imparting easily digestible consumer guides for prospective students. It has also led to the public demand for accountability from its public institutions; where before the workings of the academy were opaque and distant for the lay user and dictated entirely by specialist interests, now they appear close at hand and a part of public discussion (LATOURET, 1987).

The communications revolution has had perhaps the most pronounced impact upon production and evaluation of research of any of these factors. It is worth remembering here, the prescient words of Walter Benjamin in his diagnosis of the changing function of the work of art in the age of reproducibility. His argument followed the lines that each technological advance in distribution and replication brings about a qualitative change in its relationship to society.

By analogy, we think of how the advent of the printing press brought about the massification of information, allowing libraries to be established the world over meaning that information became much more widely available, we now enter the age of universal open access to information. The Internet has opened up a wealth of information to users larger than any library, regardless of geographical location or status. It connects academics the world over instantaneously with colleagues, and allows for collaboration and production networks that each year grow denser and more established. It allows real time and public measurement of the impact of research and even open source collaborative editing. These factors have contributed not to a new kind of research designed to replace old notions, but to signal a blurring of lines and definitions that significantly complicates attempts to find a universal measure of impact or quality, and in identifying involved stakeholders.

3.1 Mode 1 and Mode 2

	Mode 1	Mode 2
Nature of problems	Cognitive. Brought about by anomalies and problems produced by research programs.	In response to a real problem; a social or political need, or industrial application.
Model of progress/ disciplinarity	Largely homogenous, within disciplinary boundaries. Tends towards mono-disciplinary.	Largely heterogenous, varies according to individual mission of research. Tends towards multi-disciplinary.
Methods of evaluation	Production measures (number of articles), endogenous academic impact measures (citation count)	Depends upon the context of application (financial, social or academic).

Mode 1 (traditional) research seeks to solve problems within a disciplinary context. Its problems are primarily cognitive; that is to say that like Kuhnian normal science they seek to solve problems that the current paradigm is at a loss to explain (KUHN, 1962). Under Kuhn's theory, a dominant paradigm becomes the context in which research is carried out, and gives its epistemological grounds as well as the criteria for success or failure. The aim of research is to explain the empirical anomalies and problems with the paradigm, the phenomena that don't fit, in order to strengthen the theory and further embed the paradigm. Work that succeeds is invariably rewarded by the academy by way of publication and prestige or through high citation counts. In Mode 1 then, scientific progress and research is driven by the internal taxonomy of the academy and not by societal need or relevancy. When research is successful, it succeeds in either advancing the current paradigm within the discipline by solving problems and anomalies (LAUDAN, 1977).

Evaluation of contribution under Mode 1 therefore consists in homogenous growth, in terms of the number of articles published, and density of research in terms of its bibliometric impact, whether by journal impact factor or by individual citation rates. Between these two it is possible to show an institution's satisfaction of the demands placed on it by academic discourse. It is

in this way that the big three rankings predominantly evaluate research.

The defining characteristic of Mode 2 research is that it is created *in the context of application*. In this the boundaries of fundamental research, applied research and experimental development become blurred. Mode 2 knowledge is by its nature transdisciplinary, reflecting the need for a variety of perspectives in the solution of complex problems, it takes place in a variety of locations outside of the traditional ivy tower and with a variety of new social, economic and governmental actors. In Mode 1, fundamental research is carried out with no particular practical end in mind; to discover new general laws of nature, or more frequently, to solve empirical problems in existing ones without sacrificing the parsimony of the existing law. In order to use this research, it must then undergo a process of knowledge transferral; applied research.

Applied research takes the new knowledge and considers a particular situation in which it has utility. What then follows is a process of experimental design, in which the applied theory is put to use through the development of new technologies. If research is interpreted in this way, bibliometric citation is an important gauge of fundamental research's transferral to applied research. The conception of Mode 2 however, considers the entire research context; from the initial problematic, through its mode of production to its dissemination. This is not the same as applied research, which is one stage of a wider process, although the two share much common ground.

The increase in communication technology has greatly increased the number of papers produced in international and transnational collaboration. The level of international engagement present in international research points to a much greater heterogeneity in modern research, both through cultural and disciplinary context. Mode 2, because it is not driven by internal demands of disciplinary taxonomy but by an external problematic, is either interdisciplinary, made by a group formed of more than one discipline, or transdisciplinary, coming from no specific disciplinary background. This means that it uses a variety of perspectives and methodologies to analyse a problematic, in line with the 1998 UNESCO declaration;

“Special attention should be paid to higher education's role of service to

society, especially activities aimed at eliminating poverty, intolerance, violence, illiteracy, hunger, environmental degradation and disease, and to activities aiming at the development of peace, through an interdisciplinary and transdisciplinary approach.”

As the bibliometric methodology of ISI utilised in the THE and ARWU rankings and Scopus of QS are normalised by field, giving heavier weighting to citations in more archaic fields like astrophysics than to more heavily referenced ones, like clinical medicine, in an attempt to control differences in citation habits and rates between disciplines. If an increasing amount of important research is (or, perhaps, should be) heterogeneous, this presents a twofold problem for bibliometrics weighted by field. Firstly, there is the risk of categorisation mistakes of research, and gaming factors, whereby research is categorised erroneously into a smaller field either in error or in an attempt to gain greater citation rates. Secondly, current bibliometric methodology lacks the granularity to accurately describe a lot of research, as it is limited to measurement within disciplinary boundaries, or as a universal figure. It is often not clear, for truly multidisciplinary work, which discipline a piece of research should be counted in, whether it should be considered in more than one giving rise to double counting problems if we are measuring output on a larger scale than that of the individual paper, or whether bibliometric indicators are a good measure of a research team’s ability to solve a real life problem.

Earlier sociological work of Latour (1987), and Habermas in his debate with Jaspers in the higher education reform movement in the 1970s (HOHENDAHL, 2011) state that far from occurring in a political and social vacuum, the context of the laboratory (and by extension, research) is highly politicised. The reasons for pursuing, or not pursuing certain lines of enquiry is informed by a variety of environmental factors and pressures, and the effects of that research is also defined in terms of its effect on the world around it. The role of both researcher and research is determined by the network of interactions and relations to society, the economy and the scientific community. Research under Mode 2 is not just conceived of in terms of its articles, books and publications but in terms of the network of actors involved in the research process.

The density of this network defines areas of impact and importance. This relates to both networks of researchers in production of research, academic networks, and the relations between research and society. Whereas traditionally the researcher plays the part of 'spokesperson' for nature, with a one-way relationship from expertise to ignorance (LATOURE; 1987), higher levels of educational development and sophistication has brought greatly increased demand for both financial and social accountability from research.

This impulse is seen in recent discussions within UNESCO conferences and papers; universities are no longer the guardians of knowledge, and excellence in higher education should no longer be thought of as confined to within its walls. The sum of a university exists in its extensive network, one that includes not only faculty staff and the student body but extends far beyond it in the social production and distribution of knowledge. In the 21st century this is imperative; excellence that exists for itself in purely academic terms is not of value to the developing world if it is not spread, shared and used to emancipate.

This can be seen in the recent financial crises that have afflicted USP, the public concern over international performance in return for investment. It is also evident in concerns related to specific research issues, such as the demand for ecological research on deforestation, global warming, the impact of shale oil fracking and nuclear power, and social issues like food security , the drug trade and global risks derived both from multilateral networks and from local populations.

This appetite for accountability sometimes expresses itself undesirably, such as when in 2013, swept along on a wave of conservative populism, the Abbott administration in Australia promised to withdraw all funding from social sciences and humanities to concentrate resources into "areas that affect the lives of everyday Australians, like medical research". This form of complete democratisation is clearly highly inappropriate, and lay understanding of scientific processes is often far from optimal, as science and research becomes paradoxically more available and more complex and archaic (SCHWARTZMAN; 1993), but it is this tendency that forms the historical context from which rankings have arisen.

The transition to the Mode 2 paradigm from Mode 1 has not been as smooth or as general as was predicted by its authors, nor has the idea itself gained universal recognition. To understand why this is, it is worth looking first of all at the claim that Mode 2 is really nothing new, as the context of application is arguably, in some areas, more original than much Mode 1 research itself (ETZKOVITZ; 2000). The categorization and separation into disciplines of science is a Weberian turn into modernity and compartmentalization, before this the concept of science as natural philosophy predated the scientific method. In this period, research was often generated in response to some specific technical problem, using a variety of solutions. Agriculture, astronomy (with navigation as its context) and engineering are just three examples of areas were more strongly driven by external necessity than internal curiosity. With this Weberian turn, cemented disciplines, focus on pure research, problem solving and homogeneity became more a way to shut out pre-existing external forces, such as interfering benefactors, rather than the pre-existing natural condition of many sciences.

The authors themselves freely admit that Mode 2 was never intended to be an in-depth empirical study of emerging research paradigms, rather a rumination on emerging tendencies. The empirical research that has been carried out has found that disciplinary boundaries are not eroding to the extent that the authors predicted they would; despite superficial changes, most researchers are very much a product of their specializations rather than becoming free-floating thinkers (VAN LEUWEN AND TIJSEEN; 2000). Much of the reaction to Mode 2 has focused on the fact that rather than representing a dramatic departure from one epoch to another, it is rather a description of two tendencies that have always coexisted, and shift and rearrange over time depending on external exigencies and cultural tendencies, and not some inherent and permanent change in the practice of science.

It has been suggested in recent years that the correct and strongest formulation of this idea consists in a third conception, Mode 0 (SEGERSTRALE,

2000) research produced in the context of power, and so when external conditions change, the focus of science does too. Indeed, the authors themselves seem aware of this aspect of the Mode 2 conception, which they stated that despite a much more equitable distribution of knowledge throughout the world, it is still likely to increase, rather than decrease global inequality as those who are better positioned to take advantage of the expansion, the rich (be they nations, firms or individuals) will have much more opportunity to do so (NOWOTNY, 2002).

Furthermore, critics have pointed to Mode 2 being merely the post-industrial counterpart to R&D, where slowly, as it becomes more integral to the academy, research and development will eventually take over completely, and the university will be indistinguishable from the R&D department (SEGERSTRALE: 2000). From this we can conclude that the Mode 2 argument's main problem is perhaps not the concept of the research itself, but the way it is applied either deterministically or prescriptively, depending upon whether a normative or descriptive reading of the book is applied.

Mode 2 has also come up somewhat short in its attempt to apply its conclusions to all areas of research, rather than the 'new sciences' (BONACCORSI, 2001). New sciences, such as bio, nano and material sciences and engineering, along with climate issues and computer science do to some extent appear to have been birthed by a context of application rather than an endogenous disciplinary demand. Climate science, for example, has taken a common identified problem and arranged research programs around it encompassing human and physical geography, meteorology, geology, engineering, administration and political science, biology, chemistry, botany, among many others. Projects are typically organized not around cognitive problems but practical ones, by typically heterogeneous research groups. What this represents then, is not an erosion of disciplinary boundaries but a movement to form a new set of practices that exist alongside Mode 1 'normal' research.

Under the previous university administration, large amounts of money was spent on the creation of new institutes and research nuclei to increase this Mode 2 type research, such as the NAPs initiative, but results have been slower than expected. This suggests several things; firstly that this type of research cannot be imposed top down as the result of directives, that money is not necessarily the limiting factor in expanding the Mode 2 paradigm, and consequently that perhaps to a certain extent these kinds of research are best driven by the curiosity of the researcher and not by external interests. This would appear to run against several of Mode 2's tenets. The current administration instead is attempting to identify promising clusters and groups already working within the university to join together with others. In this, better and more sensitive measurement tools could greatly aid and encourage the formation of more of this type of research.

NPK advocates have also tended to overemphasise certain aspects of the work a university does, and in doing so run the risk of reducing their descriptions into one dimensional caricatures, much as rankings do. It is to be expected that the epistemological or disciplinary perspective of an author will greatly influence their vision of what a university is. Those from the humanities and social sciences will tend to emphasise the university's role as a critical body and discursive space for the improvement of society. Its views on how a university should be assessed and run are based upon this set of constituent beliefs and identities.

Researchers with background in exact and pure sciences will tend to privilege the advancement of scientific knowledge for its own sake as the ultimate ideal to which a university is held. The way in which this group would like to be assessed is markedly different to the first group.

More different still are those from engineering and new sciences, who consider a university's ability to innovate and provide technological solutions for economies. Olsen's dynamic does not just apply to how external and internal stakeholders view and benefit from universities, but also to how diverse internal stakeholders' perspectives can be. Because of this, although the literature

refers to 'universities' as one part of the relationship between industry and society, they should never be treated as monadic entities with shared interests and expectations.

Previous work on university governance has focused around four main organisational models; bureaucratic (weberian), collegial, political and symbolic (garbage can) (PUSSER, ORDORIKA; 2001). The bureaucratic analysis posits that universities share the characteristics of a Weberian bureaucracy, with fixed compartmentalisation and division of labour, and a fixed decision making process according to a hierarchical structure. This analysis comes from the Weberian division of the sciences in the 19th century, with departments strictly defining areas of study and expertise, rather than the pre-modern Royal Society model, in which scholars were expected to be conversant in all matters of science and philosophy. However, this structure can only account for formalised, hierarchical power, and not on power derived from mass movements or common consensus, and has serious problems explaining mechanisms of change, or why change occurs within institutions. It cannot account for where interests and visions of a university diverge.

The collegial model represents the self governing republic of science, with competency based upon professional expertise and placing emphasis much more on decentralised power and decision making by consent. While this analysis would probably have been insightful in the days before USP abolished its system of chairs, today it is of little descriptive value when trying to understand relationships between states, universities and outside factors, and has little insight into decision making processes if all departments are taken to be substantially autarkic.

The political structure model suggests that universities are in fact miniature political systems dominated by bargaining structures and coalition building. In comparison with the first two, it gives a stronger account of the modern university's plurality of interests and influences. However, this account tends to overemphasise conflict and cannot come up with a convincing account of institutional stability, or why by and large, systems remain relatively intact

through institutional change and conflict. As universities have proven to be remarkably stable, persistent institutions throughout history despite the absence of any single unifying aspect shared actors, it is unlikely that they can be described as purely political collections. Correspondingly, as universities have proved open to evolution with changing contexts and necessities, they would be much more constrained and less open to change if they were explainable purely as bureaucracies for organising researchers and teachers.

The final, symbolic model is also known as the garbage can model (COHEN, MARCH; 1972), which states that institutions are held together by collective beliefs and values, from which governance emerges without a necessarily set administrative or decision making process. With this image of the university, the management and articulation of all these competing ideals, interests and autonomies becomes complicated. The challenge of accommodating such an array whilst remaining in some way able to maintain a form of institutional unity means that the traditional corporate models that were rather unsuccessfully introduced during the 1980s in the West proved inflexible to the demands of an autonomous academic community whose outputs cannot always be predetermined, meaning that the corporate nature of the university tends to take over.

Likewise, the more broadly democratic models favoured in Europe and later Latin America proved insufficiently adaptable to change, as change is internally instigated through bargaining processes and not from societal exigencies, the academy tends to turn inwards and becomes in thrall to those who work within it. This means that they have an internal problem with sharing decision making and developing innovation. This is why this was broadly dropped from the agenda in Europe over the past few years, because universities tend to create an accountability gap with the rest of society under these conditions.

3.2 Triple Helix Industrial Relations

The triadic university-state-industry relations are described as a constantly moving triple helix configuration. The national systems paradigm conceives of static relations in which a university is a component in an innovation system, with fixed relations with industry and the state the act of measuring this type of system or contribution is straightforward, as results are always predictable because relations are stable. This is despite the fact that universities are complex, multifaceted and ambiguous institutions. The model assumes the engagement of only certain parts of the university, despite referring to it as a whole. If the university always plays the same role in the triad, producing patents for industrial partners, for example, then this third mission can be measured by the number of patents the university registers itself, or how much income it generates through patents produced. The triple helix formation is unstable and unpredictable (ETZKOWITZ, LEYESDORFF; 1995), it is in constant flux, in some places crystallising into stable relations, where individual models are discernible, and others where actors, definitions and behaviours are ambiguous. Because third mission (social/economic) projects within the university are generally produced in the context of their application, the configuration of relations will adapt to suit the problem at hand. This means that, as with academic evaluation, there is no single indicator to represent excellence or performance, more likely a series of partially overlapping indicators.

Any system of network analysis can be divided into components (nodes), relationships between components, and functions of a system. Generally, it is the nature of relationships that any measurement will try to represent; the density of interaction between the university and industry or the state. However it is the nature of the components and functions that will dictate the nature of the relations. Triple helix relations have been separated into three general ideal types, representative of different arrangements.

The first is the statist approach, in which the state is the key coordinating actor, and predominates over the other two components of industry and

university. This is characteristic of heavily coordinated economies, like China and some Latin American countries, generally it is regarded as a failed model of innovation system as it does not generate sufficient incentives for researchers or industry, and excessive state direction constrains the autonomy necessary for creative destruction (SCHUMPETER; 1943, ETZKOVITZ; 2002).

The second ideal type, *laissez-faire*, is characterised by very little central governmental steering and strict boundaries between components; the state, the university and private industry are strictly separated. This configuration is typical of liberal market economies, where government intervenes little in the economy. In this configuration industry is the driving dominant actor, because it supplies the funding, state and university are auxiliary support structures aimed at providing fertile conditions to profit from innovation, and skilled work force for research and development, respectively.

The third balanced conception does not have strictly defined spheres for components. Universities are no longer mono-spherical (insulated from the other two components, but rather multi-spherical (hybrid) institutions, where industry finds itself inserted into the institution, and vice versa. This type is characteristic of the transition to knowledge societies, and is the most fertile source of potential innovation at the interfaces between components. In this configuration, universities are not merely contracted by industry to produce research, they are full partners in the process, playing an active role in the shaping and directioning of the research process (LEYESDORFF, RANGA; 2013).

3.2.1 The Neo-Institutionalist Approach

Through the use of case studies, some scholars have pointed to the nature of institutions as being the key factor in determining relations. In a case study of the Caixas do Sul incubator, the authors identified the strong leadership of the Universidade de Caixas do Sul (MELLO, ALMEIDA; 2005), and its will to integrate itself further into the Brazilian innovation system as the principal driver in the formation of the triple helix relations, offering both skilled labour and the

ability to register patents for small spinoff firms. In Chile, the adoption of the Triple Helix Model has been widespread, and it is the potential social benefits that the university can offer that has been the main driver there too.

3.2.2 The Neo-Evolutionary Approach

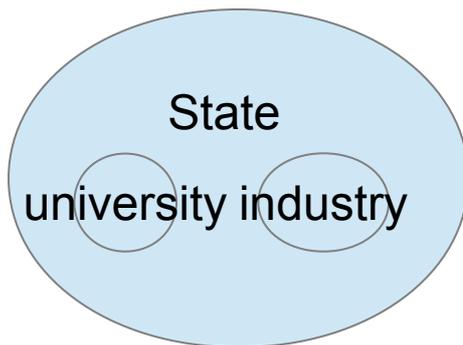
The Neo-Evolutionary approach seeks to place the growing tendency for hybridisation of universities and the convergence of industry, university and state as a result of the functions of interaction. It views each component as a coevolving subsystem in which the types of interaction are in continual change. Technological advances have greatly increased the need for much more specialised labour in the research and development process, and closer sectional integration.

The neo-evolutionary theorists are heavily influenced by a theory of social systems communication (LUHMANN; 1997), which conceives of each sphere as autopoietic, or self-creating in the way in which each receives and processes its environmental information. In contrast to the neo-institutionalists, neo-evolutionary theorists state that the nature of the institutions is constantly changed in response to the unstable nature of the relations; universities will modify their behaviour in response to new exigencies put forward by acting in incubators, technology parks and in the process of commercialising research.

The way in which an academic approaches a problem is defined by the cultural aspects of academia; of academic rigour, autonomy and empirical method. Industry and states, on the other hand, are defined by different sets of values and epistemic perspectives, applied to the specific problem at hand. The neo-evolutionary perspective conceives of a set of co-evolving subsystems modulated by an overlay of recursive networks, which constantly change the nature of each component. These networks take the form of discursive spaces, of technological changes and new hybrid forms of interaction between universities, the private sector and the public sector. This means that in the case of an incubator, where public policy, higher education and the private sector combine, it is not necessarily that all three converge with their own systemic perspective and fixed role within the partnership, but rather through

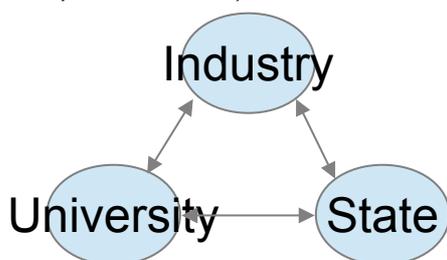
this sharing of space, all three adapt norms to the new discursive space, bringing new requirements for leadership and management, they create hybrid institutions of all three. This scenario is one of 'endless transition', rather than an 'endless frontier' of basic research in academia.

Fig 1. Statist Conception (Coordinated Market Economies: China, Latin America, Eastern Europe)



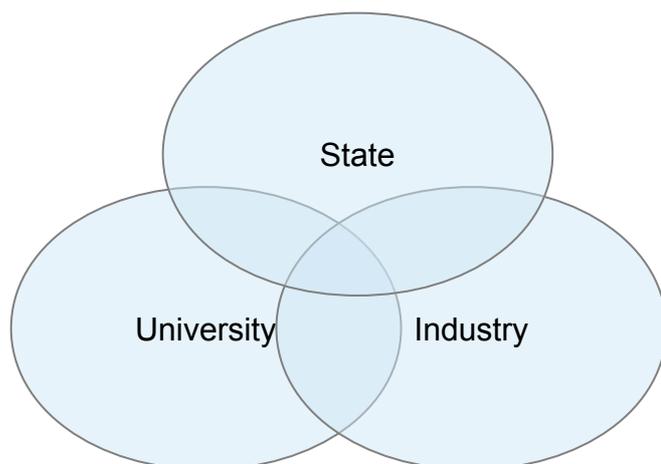
In this configuration, all relations are mediated and controlled by the state. This means that university and industry interact directly relatively little, and usually geared towards some state goal. Funding comes from the state into the university, which in return is expected to generate economic development through innovation. On this model we would expect to see a high number of patents produced by universities and in industrial co-partnerships, and relatively few co-authored publications, because firms will be relatively less likely to wish to invest in research, as the state will generally take on this role. Universities are more subject to the 'Instrument of state goals' vision, and therefore less reliant on production oriented to the demands of industry as it is less dependent on private money.

Fig 2. Laissez-faire arrangement (typical of Liberal Market Economies: USA, UK, some Western European countries)



In the laissez-faire configuration, the state provides (some) funding to universities in order to form a productive workforce for industry, the university provides, in turn, knowledge and skilled labour to industry. Industry then produces innovation through its research development programs, generating income which it then feeds back into both higher education and the state in the form of investment and taxes, respectively. A laissez-faire environment will produce high numbers of co-authored publications, but universities will produce little, if any, intellectual property in the form of patents, as industry will naturally seek to maximise its own profits by retaining the profitable part of the research process (the patent). The laissez-faire model fits the vision of the university as a tool in service to a market economy, as industry is the main driving force of this configuration, universities are more dependent on private money and therefore more oriented to the interests of a market economy.

Fig. 3 Balanced configuration (Transition to a knowledge society)



The balanced model sees leadership based upon the nature of the problem, with the university taking a more active role in processes in many instances. Flows of capital, resources and influence are much more ambiguously distributed throughout the triad, and subject to constant change. What this means is that neither the state, nor the private sector assumes leadership over projects, rather that the configuration is dictated by the specific needs of the project at hand. This mandates the formation of hybrid institutions, rather than single sphere institutions, as we see in configurations 1 and 2. While a single sphere institution is highly centralised, hierarchical and characterised by relatively little interaction with other spheres, a hybrid institution is one that

operates at the interfaces between institutions, typically with very high levels of interaction, a much smaller and more streamlined hierarchy that shows a reflexiveness and ability to change according to conditions. It should be noted that in the literature the term institution is used somewhat liberally, in recognition of the fact that each individual sphere has a different organisational structure and a different definition of institution (a university is not the same as a firm, and neither is the same as a government department). Because of this, an institution is defined in terms of its conceptual space. Therefore, when we talk about hybrid institutions we do not mean an entire university as a hybrid institution, but a technology transfer office, an incubator, a technology park or even an individual collaborative project. This change in definition of institution allows easier exchange between the spheres.

Where classic innovation systems conceive of industrial relations as a process involving only knowledge transfer between spheres, typically from academia to government or industry (RANGA; ETZKOWITZ, 2013). In hybrid institutions, this is extended to full collaboration, networking, substitution (where a deficit in one sphere can be taken up in another) and collaborative leadership. This closeness reduces the potential for conflicts of interest related to generating profits and use of manpower, and encourages a more mutually beneficial 'win-win' scenario. This drives innovation away from the low-risk, low-gain model where individual institutions do not wish to risk large amounts of resources into projects with a risk of failure by sharing burdens and risks with others, allowing much higher risk, potentially much higher gain projects, greatly accelerating technological development.

Etzkowitz conceives of this leadership as depending on what he calls an 'innovation organiser', someone, or a group who can come from any of the three spheres, or a combination of them, with the necessary drive, support and organisational capacity to drive the project forward. In this way, the balanced conception relies not so much on investment or financial power, as the statist and laissez-faire conceptions do, but on individuals and ideas.

Brazil's regions are highly heterogynous in terms of development and funding investment climates. In most of the country private investment in R&D is extremely low compared to state investment, despite the Technological Innovation Law (LEI N^o10.973, DE 2 DE DEZEMBRO DE 2004, chapters II and III) giving specific and institutional framework and support to collaborative hybrid processes and support from FINEP. This contributes to a broadly statist model of innovation as firms are not active enough partners in the innovation process. In the state of São Paulo private investment outstrips public investment in research and development, and has increased steadily in the years since the law's introduction. Because of this, São Paulo sits much closer to a balanced conception than the rest of Brazil, despite being under the auspices of the same legal framework. This is because of the range of other factors at play in the state and absent from others; the higher concentration of high-tech industry, universities and qualified labour, for example.

3.2.3 Consequences for measurement

In the statist and laissez-faire conceptions there is a clearly defined set of inputs and outputs, and so measurement is relatively easy to determine. Measurement of a triple helix system must find a way to be sufficiently adaptable to represent the constantly changing relations between the private sphere and the public. In this we find the RUF metric of number of patents granted to the university to be a very limited snapshot of one specific type of activity, assuming that the university is not transferring knowledge to create innovation, but rather producing innovation itself. On this metric, we can see from the benchmark project in the appendix that the University of California Berkley, which has very strong industrial ties, especially when compared to Brazilian higher education, would score zero on a metric of this type because it does not register its own patents, but produces applied research in conjunction with industrial partners, who then produce the patents themselves. We should not assume uniformity across different systems, institutions or even

agreements. In recognition of this fact, the Times Higher measures private money attracted by the university each year to represent industrial insertion.

USP does not at present measure all private money entering the university, and it is not a reported score for any of the U-Multirank benchmarked universities in the benchmark project. The way in which this data is recorded is not clear for the THE, if the data is not accessible to the university itself, the possibility that it is available to an external ranking agency seems unlikely.

There is a problem, furthermore, with the use of private money attracted as a normative metric with which to rank institutions. Money attracted is not an indication of performance or return on investment, nor should it be regarded as an unconditional good, like public money is (SLAUGHTER; 2001). Industrial resources attracted is an institutional input, it is only loosely related to performance; companies will be more likely to invest if they see the university as being a favourable development environment. However, they may not invest because they see the university as the best, rather the most amenable to their needs. Private money can be used to distort academic incentives, constrain autonomy by dictating research programs oriented towards 'low hanging fruit' and short term investment, it can be a strong indicator of the decline of the availability of public funding, or a sign of a university's inability to attract it. The entrepreneurial university measured like this provides a dangerous incentive towards states to reduce public funding for higher education if institutions that attract more private than public money are normatively better, more productive and more competitive.

Economic success and innovation makes up an increasingly important part of a university's mission in the twenty first century, both because of the recognition on the part of universities of their social role in fostering the production of a knowledge society, and in the internally generated demands of modern research. It cannot predominate over the other missions and functions of a research institution. If higher education becomes dependent on industrial

investment to fund research, this will happen, as the logic of research and development favours short term profit over long term discovery, profitability over social benefit. This means that if financial success is measured purely in these terms, in a normative way, the incentives generated will be towards what is profitable rather than what is scientifically or socially important.

3.2.4 Unified measures of quality

This conception of modern research points to serious consequences for the possibility of constructing universal evaluation; if research is produced within the context of application, it is its response to this context that forms the measure of impact. This means that there is no universal measure of impact, nor of quality. Instead we look to a conjunction of partial indicators (MARTIN; IRVINE, 1980), many indicators explain some aspect or facet of quality, no single quantitative measure is the key universal. The consequence for ranking exercises is serious; if we can only point to a matrix of partially overlapping indicators of quality, then any ranking can show at best a partial picture.

Recent years have seen the rebirth of qualitative public evaluation exercises, such as the Research Assessment Exercise in the United Kingdom, and the Research Quality Framework (RQF) in Australia, using citation information as a guide but considering a broader range of factors. Response to these has been extremely mixed, with many researchers concerned about feeling pressure to make their research 'accessible' and 'marketable' as opposed to innovative or useful, but they are indicative of a broader trend within policymaking away from such blunt measures of impact. Rankings, however, do not incorporate such considerations as a result of the difficulty of incorporating qualitative judgments of impact into quantitative studies.

The original ISI index created by Eugene Garfield in 1963 was designed to describe this academic network and assess variations in publication patterns across national and disciplinary cultures. For this, it still plays an invaluable part of any understanding of academic impact, however its use in composite rankings as the only measure of impact begins to distort publication behaviour by only rewarding certain kinds of publication. Furthermore, although it can

consider variations in average citation rates across disciplines, it has a major problem with research that does not fit into any one discipline. In this way existing bibliometric methodologies can be a very good measure of endogenous, Mode 1 type research, but not for Mode 2, heterogeneous growth of research.

Until very recently, the only way in which impact was quantitatively assessable in this way was through citation counts, but with the rise of the internet we now have the ability to conceive of dispersion and consumption in a much more holistic manner, with the ability to measure internet 'traffic' through hits and downloads, and social network analysis (Hoffmann;2014) which can give a measure of research's social impact, measure spread and citation density rather than simple counts through the use of Eigenfactors, and map research around topics, tags and broad themes that interest a variety of disciplinary backgrounds (Bornmann;2014). This means that today, journal impact factor (JIF) and citation rates are just one among many measures of research impact. Their use as the only one suggests that many rankings prefer to measure what is convenient to measure rather than what is actually informative (Ordorika; 2008).

3.3 Knowledge as a Global Public Good and the Open Access Ecology

Due to the centrality of knowledge in the role of development and innovation, knowledge, and by extension research has the character of a global public good (STIGLITZ, 1999). Global public goods have two key properties; non-exclusivity and non-rivalrous consumption. Something is a global public good if its consumption does not necessarily deprive another of the opportunity to consume it, and nobody can be denied access to it. Learning and utilising a scientific theory does not deprive anyone else of the ability to also use it, and, within certain constraints (ability to understand it, access to education) nobody can be, or should be, deprived of the right to access it (UNESCO, 1998). Non-rivalrous consumption means that there is no marginal cost associated with sharing knowledge with others, often it is beneficial to have the input of outside sources in development. As with other global public goods, the marginal costs associated with knowledge sharing consist in the cost of transmission; in

teaching, and in publication, not in the sharing of the knowledge itself.

From the condition of non-rivalrousness, it then must follow that nobody can be excluded from consumption of knowledge. This means that knowledge cannot be privately supplied; if knowledge is public then it is impossible to profit from it, as everyone has access to it then competition drives its price towards zero. Where knowledge has application to a practical end, and considering that there is no marginal cost implicit in knowledge sharing, this knowledge should be made as freely available as possible. This means that any unnecessary restriction on the spread and sharing of knowledge is retards the process of development and slows the spread and evolution of knowledge.

If knowledge is not a private, transactional good, we can see it as a post-capitalist good, something which feeds and lives alongside capitalist goods and innovation, but is not itself a transactional good. This means that the market absorbs and depends upon research, either directly through the intellectual property regime or indirectly by creating fertile conditions to innovate and share information. In this way we see that the relationship between the economy and knowledge are much more complex than might be commonly supposed rather than being one and the same thing, as is often supposed by neo-liberal theorists, and a good proportion of NPK theorists (BARTUNEK, 2011), who lack the ability to differentiate in the context of application between commercial and social contexts, instead seeing them as bundled. Instead, the relationship between economy and knowledge economy is mediated by society; by ideology, politics and most crucially status (HAZELKORN, 2013; MARGINSON, 2009; ORDORIKA, 2010).

In practise, this normative conception often fails to hold. Privately owned knowledge, such as that protected by intellectual property law provides, in a constrained way, the exclusive right for private individuals to profit, for a limited time, from the fruits of their research. Often military, defence and nuclear research is held from common consumption by governments as a matter of national interest, while competing laboratories and scientists may withhold information from one another, albeit usually temporarily before publication (LATOURET, 1987). Furthermore, the publishing industry as it exists today does not serve this purpose, creating artificial scarcity in a resource theoretically

infinite (WELLEN, 2013; 2004). The commodification of knowledge occurring both through the growth of the private sector of higher education, and the liberalisation of the research industry in the US under Reagan means that the model perpetuated by rankings is that of the old elitist order (MARGINSON, 2009). This is despite the incipient Open Access movement, propelled both by public funding bodies, multilateral organisations and institutions themselves.

As the marginal costs of publication fall with advancing technology, eliminating both rivalrousness in consumption (my use of a pdf does not inhibit any other user from having the same file) there is both a scientific and social incentive to ensure that that access to the information produced is kept as low as possible. At present, this is not happening with the huge paywalls imposed by organisations like Thomson Reuters, who have preferred to pass the benefits on to shareholders rather than to users.

3.3.1 Green and Gold Open Access

Following the imperative to open access to information, there are at present two commonly used routes to open access, green and gold (BJÖRK et al., 2014; HARNAD et al., 2008). Green access is straight-to-press open access via either an open access journal or a database such as ArXiv, SSRN or a university repository, while gold, an increasingly common requirement of funding bodies, is the release of previously commercially published work after a given grace period, usually 36 months.

The impact of the appearance of open access journals has been enormous, and is growing constantly, especially in the developing world and institutions outside of the educational centre. It is contributing to a significant democratisation of world research both in terms of access to publishing for those without contacts and prestige in the more traditional academic centres, and in access to information, challenging the conception that published research is a commodity or intellectual property. In line with Latin America's commitment to higher education as a public good, open access is the single

best way to connect centres of research excellence to populations, combining the second (research) and third (extension) pillars of higher education.

That is not to say that there are no detrimental effects from this levelling of the playing field. There are growing reports of exploitative predatory publishers accepting large figures from individual academics to publish in fraudulent journals, and several worrying scandals related to falsified and plagiarized works being published, now that in many cases the process of peer review has been significantly reduced or altogether eradicated. This signals that probably in the future academics will need to be more careful and critical in citing work, or that publishers will have to come up with more novel ways of replacing the expensive and slow peer review process. In some cases this is already happening.

4. RESEARCH AND IMPACT

The main benefit that the Universidade de São Paulo can gain from aligning more closely with international metrics is in the field of impact of research. It is in this area that the university struggles most in Shanghai Jiao Tong, and in the movement from an elite system to a mass system it is time for the focus to move from raw research output, in which it is extremely competitive, to the actual effect the research produced has, which is still appears by most existing measures to be lagging well behind.

However, measuring 'real' research impact, the full effect of research, is one of the most contentious and difficult things to measure in higher education. Bibliometric impact is usually used as a proxy (GLÄNZEL, 2003; LETA; PEREIRA; CHAIMOVICH, 2005; SEGLEN, 1997).

Drawing from the UNESCO declaration, there are three main relevant ways of thinking about research impact, although the three share significant crossovers. Academic impact is the measure of the density of connections generated within the academic community, the extent to which research contributes to the furthering of scientific knowledge and contribution to global academic discourse. Social impact can be seen as the extent to which research affects the lives of citizens outside of academia, whether it is in influencing public policy, contributing to public education or improving the well being and

quality of life of citizens. Economic impact is the extent to which research generated translates into an industrial or technological setting, and could be measured in terms of income generated and patents derived from research.

There is significant scope for crossover between these definitions; a highly cited piece of fundamental research may then generate patents and industrial utility, while something that has a high level of social impact may have an impressive academic impact as a result. Similarly something that has a high level of economic impact may create innovations that have consequent social impacts. However, there are circumstances where research performs well in one and not in the other two, and there are circumstances in which one or more of the types of impact are simply irrelevant. Theoretical physics, for example, is likely to have limited social impact, while the industrial application of many social sciences is likely to be negligible.

4.1 Academic Impact: Journal Impact Factor and Citation Impact

Academic impact is the extent to which a piece of research, or a researcher influences the global debate on the topic to which it refers. Rankings take this aspect into account almost exclusively, as it is the easiest to measure of the three, its boundaries relatively well defined.

Citation impact, as measured by Thomson Reuters ISI and Elsevier Scopus, is the number of citations an article published and indexed on one of their networks receives from another article also indexed on the same network. The assumption is that the higher the number of citations, the greater the impact, and therefore the higher quality a piece of work is. Citation impact is then normalised by field, according to the 21 different categories and 554 subfields defined by Elsevier and Thomson, to control for disciplinary variation. This is considered the most accurate measure of an article's impact, and in many cases would appear intuitive; if article A is cited more times by other articles than article B, the amount of debate it has stimulated is higher and therefore its impact is higher. This argument is predicated on the assumption that all citations have the same, or similar value. In reality, the reasons to cite an article could be, among many others:

- To form the basis of an argument.
- To make reference to a relatively inconsequential part of research.
- To cite precedent.
- To highlight a poor argument.
- To reconfirm previously carried out research.
- To refute previously carried out research.

In other words, citation count can ascribe no particular value judgment to an article, sometimes it is because it is of central importance to the argument, sometimes of only marginal importance, sometimes it is a positive endorsement, while other times it is to highlight or refute poor scholarship. Indeed, some articles are cited many times because they are exposed as poor scholarship, not because they are of high quality.

Similarly, we can identify many asymmetries on the part of where the citation is made. If we consider that academic communication takes the form of a network, then citation count only takes individual links into account, and not the network as a whole, or the impact upon the network. If an article is cited by another which is itself highly cited, then this citation can be said to have significantly more impact than a citation from an uncited paper. Simple citation count cannot account for this, which is why tools like the Eigenfactor compound indicator have been developed, and increasingly network analysis is used in the field of scientometrics, rather than citation analysis.

The alternative to citation count often used is the journal impact factor, in which the citation impact of the article itself is not counted, but represented by a mean score of citations per article published in a given journal for the year. Like the Highly Cited list, the Journal Citation Report (JCR) is published annually by ISI, and used by the Times Higher Ranking because its data are much easier to gather and use than the more complex citation count. This means that what is in fact measured is not individual article impact at all, but a reflection of the prestige of the journal in which it is published.

JIF is easy to manipulate for unscrupulous editors, with some reports demonstrating that journals have managed to boost scores by up to 10% by changing the formatting of the journal. The denominator in the JIF calculations only takes into account articles, it does not consider letters to the editor, reviews and other marginalia, whereas the number of citations, or the numerator, does (TOGIA; TSIGILIS, 2006).

4.1.1 Variation between Disciplines

The criticism of interdisciplinary comparison of citation or journal rankings is well documented, and centres around two main claims; first, variation in citation rates, second, variation in publishing cultures and habits.

Of the three prevalent measures of impact factor; impact, immediacy and half-life, there are vast differences in citation cultures between disciplines. First of all there is the issue of the relative size of disciplines. It is an untrivial fact that some disciplines are simply larger than others; there are more active

researchers within them, and higher numbers of papers published each year. Within this context a paper or journal that is moderately well received in a field like clinical medicine will be cited many times more than a top 1% paper published in a more archaic field like astrophysics.

This is controlled for by Thomson ISI and Scopus by adjusting for field or subject size by giving relative scores, so is possibly not as large a problem as it might be. However, adjusting simply for field size will not account for everything. In many scientific disciplines it is customary to reference huge numbers of papers in research, citing as many previous precedents as possible, while often in social sciences the number of citations per paper are often much smaller, and only directly related to the part of the argument in question. In this respect citation density plays an important role in the total score, and while at the moment field scores are normalised by total article number, they are not normalised by citations made per article. This creates an inherent bias when aggregated towards institutions that are exact science-intensive.

Also of concern is that citation rates vary between disciplines. While natural sciences tend to have much higher immediacy rates as the result of the need for other researchers to have up to date information, they also tend to have much shorter half-lives, as natural sciences *tend* to follow a linear model of progression. This means that scientific papers from twenty or thirty years ago are referenced much more infrequently, as the problems they answer will have been repeated and refined by subsequent work, their problem solving tools improved.

This is both predicted by philosophy of scientific programs (KUHN, 1962; LAKATOS, 1969) and borne out by empirical observation. The social sciences however, cannot be defined in such terms, as progress in social sciences and humanities, if we may use such a term, is non-linear. The role of historical debate and controversy plays a much more significant role, with theorists often drawing on arguments from the past to current debate in a way that is not present in natural sciences.

In Popperian terms, (POPPER, 1959) social sciences are pseudo sciences; they fail the falsification test; Kant still forms a significant part of

humanities reading lists, phlogiston biology does not form the basis of medical science reading lists. Coupled with this much longer half-life, there is also a much smaller immediacy rate in most humanities and social sciences, as new ideas take much more time to percolate through academia. Therefore, influence in social sciences is often much harder to distinguish within the 5 year time frame measured by the rankings. Again giving privilege to a small tranche of natural sciences.

At the other end of this temporal horizon we find some of the new technological and interdisciplinary fields, like computer science, biotechnology and biochemistry for whom immediacy rates are so fast that the disciplines have found that by and large the refereed journal format is too slow for the efficient dissemination of science; often the lag between what happens in the laboratory or conference room and what is published means that what is eventually published is superseded before it even comes to press. In part, this is a result of the administrative 'traffic jam' brought about by the peer review system necessary for publishing in a top ranked journal (BREMBS; BUTTON; MUNAFÒ, 2013). For these disciplines, freely published conference proceedings and green route open access are the main methods of scientific communication, and much published output takes the form of reviews and meta-analysis. In this case, rather than being too short to fully capture impact in a discipline, the way the rankings structure their data is not quick enough to capture it.

The use of this information holds the implicit assumption that refereed journal publishing is the predominant channel of communication for all disciplines. This is not always the case, as with the disciplines outlined above, who prefer quicker methods to publish their work, and in the social sciences and especially humanities, where book publishing and monographs play a central role in dissemination. ISI or Scopus does not measure these, and therefore for a number of disciplines these rankings fail to capture the bulk of academic communication as a result of using a template taken from a few traditional natural sciences, for which it works reasonably well, and applying it universally to academia.

4.1.2 Bias against multidisciplinary research

In accordance with article 6 of *Higher Education in the Twenty First Century: Vision and Action* (UNESCO, 1998) and the NPK literature on modern research tendencies (NOWOTNY; SCOTT; GIBBONS, 2003a), inter- and multi-disciplinary research plays a much more central role in the research landscape than ever before, with the expansion of research collaboration and applied research fostering a diversified approach to problem solving than in previous eras. It would seem clear then, that any form of measurement of research quality should take this movement into account.

Measuring by journal impact or citation impact, however, appears to be prejudiced against multidisciplinary research (RAFOLS et al., 2012). As a result of the definitions used for categorising research along disciplinary lines to control for variation between disciplines, there is then a problem with categorising research that either crosses disciplinary boundaries (multidisciplinary) or that does not fit easily into any disciplinary field (transdisciplinary) (LEYDESDORFF, 2009; OPTHOF; LEYDESDORFF, 2010). While there is no significant variation in citation rates between mono and multi disciplinarity in the social sciences, average citation rates are often much lower for interdisciplinary research in the sciences. This is probably a result of the much more relaxed approach taken within social sciences to disciplinary boundaries, meaning authors will cite more frequently from outside their own fields. However, as rankings take a disproportionate weighting from natural and life sciences, this amounts to a bias against multidisciplinary research as a result (LEVITT; THELWALL, 2008).

Alternatively, for a select group of multi-disciplinary research, the rewards in terms citations can be very high. This is often for articles published in some of the highest cited journals on the index, such as *Nature*, *Science* or *PLOS One*. In order to be published in any of these, the article cannot be too specific to any discipline but of general scientific interest. The potential for multidisciplinary research to have wider ranging effects over academic discourse are clear, however in broad terms, we can see multi-disciplinary research as higher risk, higher reward research.

The field 'interdisciplinary' used by Thomson ISI is clearly too heterogeneous to qualify as a unified group, serving more as a miscellaneous category for research that does not fit into others, meaning that analysis is left using old disciplinary notions to force research into categories in which they do not easily fit (LEYDESDORFF, 2009).

4.1.3 Variation within disciplines

As if to confound matters, it is not even clear that we may use impact factors to represent impact within the same specialised field, as this promotes the continuation of orthodoxies and mainstream research at the cost of less considered areas. For example, in FMUSP's hepatology department, itself a sub-department of gastroenterology, Dr. Alberto Farias is widely published and well respected. Over the past few years he has published on both alcoholic cirrhosis and schistosomiasis. His papers on cirrhosis are well received, and as a result are highly cited. It is in schistosomiasis research, however, that Dr. Farias is renowned. The citation and uptake by the wider community however, is much lower than for research on cirrhosis. The reason for this discrepancy is that cirrhosis is a prevalent condition in developed countries related to obesity and alcohol abuse, and therefore composes a significant percentage of the centre's research in hepatology. Schistosomiasis, on the other hand, is a disease transmitted by parasites on snails prevalent, and in some areas endemic, in agricultural areas in the north of Brazil, sub-Saharan Africa and Southeast Asia. Its geographical distribution has a direct correlation with the amount of scholarly attention it receives, and as a result is categorised as a neglected disease by the WHO.

If we were to make value judgments on the two based solely upon citation rates we might, in this instance, be led to some counterintuitive conclusions, and therefore some undesirable decisions when considering allocation of funding resources. FMUSP's relative contribution to the area of schistosomiasis research is much bigger, while its absolute contribution to research on cirrhosis is greater. We cannot decide on the basis of citation, or journal impact factor, which is more important or of better quality.

This problem is further complicated when we turn to the social sciences, which have a tendency to be much more heterogeneous, sharing fewer fundamental epistemological positions than do natural sciences. This is especially clear in divergent academic traditions between continental Europe and its Anglo-American counterparts, in analytic and continental philosophy, or in clinical psychology and psychoanalysis, for example. The imposition of citation metrics pre-assumes a universal meta-language of research, de facto legitimating mainstream, usually Anglo-American orthodoxy at the expense of field diversity or divergent perspectives (ERNE, 2007).

Because of this, we can consider that impact factor is a very useful descriptive element in any judgment of research quality. The reason we know so much about publishing habits is through bibliometric analysis, and in this respect provides invaluable descriptive insights into global research behaviour. What it is not, however, is a universal indicator as the rankings use it. In the field of bibliometrics it is generally agreed that significant comparison can be made only when the two articles to be compared are within the same discipline, the same publishing culture, and published at similar times. Even then the intruding issue of topic choice may inhibit valid comparison; comparing two articles on cirrhosis is appropriate, comparing cirrhosis to schistosomiasis on the basis of aggregate citations is not. This is a level of detail not recorded by Thomson ISI or Scopus, because breakdown by topic would present a potentially infinite list of categories, if we were to look at the whole of academia.

Impact factor is a partial indicator, which needs to be supplemented by further information in order to give a clearer picture of impact, or of quality. Its use as a universal places unacceptably excessive weighting on one aspect of quality, which in turn has a distortive effects on publication incentives. While it is true that highly cited papers and researchers will usually perform well in other metrics, therefore justifying the use of impact factor as a proxy for quality, there is much evidence to suggest that in this case the reactivity, the way in which the phenomenon measured reacts to measurement, is not healthy (NELSON; SAUDER; ESPELAND, 2014).

4.1.4 The Geographical Problem

If we are to use citation rates as a way of measuring performance, we also need to establish that they are indicative of individual institutional performance, and not a reflection of geographical location or a reflection of environmental factors. Using Altbach's centre-periphery argument, we would expect that the centre, characterised by dense networks of collaboration and citation would be more heterogenous, and better performance would indicate a university better able to take advantage of this, while the periphery, characterised by much weaker linkages and networks, will be less able to do this. For this exercise, U-Multirank was used to take the data from three distinct groups of 'periphery' universities; Latin American flagship universities (both private and public), non-English speaking leading Asian universities, and finally Anglophone leading Asian universities. They were then grouped accordingly. The difference between Latin American institutions and non-English speaking Asian ones

is statistically significant ($p < 0.001$ on an unpaired t-test):

USP	0.643	TSINGHUA	1.097	HKU	1.165
UNAM	0.611	PEKING	1.057	HKST	1.318
UNICAMP	0.662	TOKYO	1.04	NANYANG	1.439
UBA	0.713	SHANGHAI JT	0.902	NUS	1.347
PUC-CHILE	0.685	FUDAN	1.093	CITY HK	1.284
UNIV-CHILE	0.688	OSAKA	0.898	HK BAPTIST	1.24
UFRJ	0.68	SEOUL NAT	0.91		
UFRS	0.663	CHINESE HK	1.163		
UFMG	0.601	KOREA NATIONAL	0.8		

This suggests several things. Because the other universities do not form such a tightly heterogenous group in terms of governance, financial means or concentration of talent, and yet display such uniform characteristics, geographical location, or “centrality” to the network of global research may play a larger part in analysis of citation rate than any notion of research excellence or internal governance factors. While it may account for differences within the groups, access to the wider world of academia based on location may be a much more defining factor in determining how well cited a university's work is on a global scale. Given that USP cannot decide to move to Europe or China, it is

perhaps limited in the amount of progress that can be made in this metric by environmental factors.

The exceptions to this when we introduce a third group to our analysis; Singaporean and Anglophone universities in Hong Kong show a statistically significant ($p < 0.0001$) difference in citation rate from its non-English speaking neighbours. This group is a mixture of public and private, and high profile and less well known institutions, all of whom show a significantly higher citation rate than their non-english speaking continental counterparts. As most of them also follow recognisably Americanised governance structures, it could be argued by proponents of the model that it demonstrates that it produces better results.

Alternatively, where English is the main language of instruction and publication, and governance structures are sufficiently recognisable to the US, universities are more inclined to cooperate with them, and cite their work. These variables merit further research, and could be illuminating on the issue of whether USP's problem is in fact producing quality research, having a governance structure that is favourable to the production of high quality research, or is in fact simply a reflection of being a Latin American university and therefore further out to the margins of world higher education than Europe or Asia. These changes are highlighted especially by the fact that U-multirank does not record its scores in ordinal scores, or in z-scores weighted towards the top ranked institution, but in actual figures.

It has often been said that because the deck is loaded against universities in emerging economies and outside of the educational centre, this should somehow be compensated for, or totally ignored by those wishing to look at them. Similarly, because of the linguistic imperialism argument (PHILLIPSON, 1992), more importance and prominence must be given to work published in languages other than English. For the long term prosperity of global research, I do not think that this is a helpful position to take. The intention should not be to insulate systems from comparison or criticism, but rather gain a better understanding of how they work. It is clear that for all but a few areas in social sciences and humanities specifically connected to local issues that in

order to contribute to global conversations on science, researchers must publish in English. It is commonplace even in Brazil for researchers to refer to the 'the English language literature' in articles, for if work is published in Portuguese, Mandarin, Farsi or Korean it is almost by design never to be referenced by international readers.

This imperative to publish in English is perhaps the best way to start to bridge the geographical citation gap. The supposition made by Phillipson is that the periphery writes in English *for* the centre, because the centre demands it. It is far from clear that this is true, however. Within the Ibero-American sphere, Spanish is often, but not always used as the lingua franca, but if we wish to communicate with Sub-Saharan non-Lusophone Africa, MENA or Asia (all priority areas of collaboration for Brazil), English is likely to be the only shared language. If we wish to overcome geopolitical handicaps and break out of having such heavily localised research networks, the linguistic aspect must be tackled.

When comparing within systems, for example comparing American institutions with other American institutions citation rates may still be indicative of individual achievement, but for the developing world, or the educational periphery, they are artefacts of a university's position in the network, not reflections of their individual contributions. This is not to say that the exercise is futile or incoherent, but that we must be more aware of this difference when we are considering citation rates, not necessarily attributing too much importance to them in isolation. If citation rates average out across institutions to such an extent, this suggests that analysis must be moved from the institution as a unit of analysis in one of two, complementary directions; towards the macro-level or micro-level.

The macro-level would be the systems health benchmarking approach advocated by the World Bank (SALMI; 2012), which seeks to analyse both nations and regions for the strength of an entire system, ostensibly to allow for greater differentiation of goals within a system, and to deter governments from ploughing money into one potential world-class institution for international

prestige at the expense of a broader system that satisfies the demands of its economy, the 'perverse incentives' rankings propagate (ORDORIKA; 2012). If citation rates average out to such an extent according to region, it suggests that building one world-class institution may be a misguided endeavour anyway. It is difficult to believe that all Latin American institutions, within a variety of governmental frameworks, legal statuses, investment levels and concentration of talent are producing near-identical quality of output, rather than citation impact being a poor indicator of institutional quality, for the developing world at least.

On a macro-level, benchmarking could begin to understand better how to close this gap, given the global distribution of research capabilities it cannot be explained purely in terms of the number of active researchers but in terms of organisation and means of communication.

4.2 Social and Economic Impact

Social impact is a broad term generally applied in Anglo-American HE to non-academic (third mission) impact. Because of this, many universities today conflate economic impact with social, however for Brazilian universities the distinction is clear; public universities are still tied to a notion of social service to the nation and local population. For the purposes of this, taken from the new production of knowledge arguments we can separate two distinct forms of social impact; one at the entry of new stakeholders and actors in a more reflexive production setting, and the other at the point of dispersal and divulgation of research, both increasing public and professional education and influencing policy decisions in the wider world.

Universities have an almost unique advantage in the production of knowledge in their autonomy from political influences to influence policy at a local, national and global level. Research produced in universities strongly influences debates on climate change, public health and social and economic development, for example. In this, impact exists not only in terms of the debate generated in academia, but the real effects it has outside of it. This impact cannot, and should not, be measured solely in terms of endogenous academic

recognition but in terms of research cited in international organization policy reports, participation in international public events, civil society organization (CSO) reports, government proposals and legislation and professional materials and clinical guidelines. In this we can measure inputs, such as number and intensity of external partnerships forged by the university, research projects undertaken, public funding captured for the treatment of complex problems, and outputs, such as citation in published external reports and government papers and social projects delivered.

Universities also have an important role to play in professional education, especially in the practice of medicine and engineering. Professionals in areas that are subject to consistent innovation and evolution in practice as research develops have a constant need for up-to-date information. The publishing of clinical medicine articles has a double purpose; firstly to inform and structure debates for future researchers, and secondly to inform doctors and nurses to ensure that the best possible care is being delivered to the operating table or clinic. Because of this, clinical medicine journals have been very early adopters of new possibilities in research measurement like Altmetrics, in recognition of the fact that, at least in this discipline, measurement of social impact can be as important as academic.

Economic impact applies to the functionality of a university within a given innovation system. It is the contribution a university makes to a knowledge economy, for the reasons set out in previous chapters; the key role universities take in determining economic growth through innovation, the need to diversify institutional revenue streams, and the need to tie research to relevant issues. The form this takes, however, is highly dependent on the particular innovation system in which a university finds itself, also increasingly to how it inserts itself into international economies. In this it is difficult to actually measure globally on an institutional level because it assumes uniformity of national systems and behavior, when as we saw in the section on triple helix relations, there are many different possible configurations of this relation, depending on environmental factors and situations.

5. THE SHANGHAI JIAO TONG ARWU RANKING

In 1998 the Chinese premier announced the intention to build several World Class Universities in China, and started the 985 Project to modernise Chinese higher education (Liu, 2013). The lack of clarity over the definition of a World Class university, or the key determinants of them, led to an attempt to quantify and measure excellence in order to form strategic goals for Chinese institutions. The Shanghai Jiao Tong ARWU ranking was launched in 2003 following extensive consultation from China's Ministry of Education, it was initially intended as an internal document as part of a wider, but strictly national project (ACADEMIC RANKINGS OF WORLD UNIVERSITIES, 2012). However, the interest, both public and academic, that it generated meant that almost by accident it became the first truly global ranking. Put together by a team at the eponymous university, its aim was to determine how China's scientific production measured up against the West's. Because of this, the ARWU is focussed exclusively on academic output and scientific production.

It is based purely in publicly available and externally verifiable information, and so avoids many of the criticisms levelled at the Times Higher and Quaquarelli Symonds about the lack of transparency and clarity in methodology. It does not contain a reputational survey, and so cannot be accused of a lack of objectivity, and is an institutional project and so is not driven by commercial interests. In this way the Shanghai Jiao Tong is considered perhaps the most rigorous and objective of the global rankings. It also avoids Ordorika and Lloyd (2013)'s criticism over volatility in rankings by only giving hierarchical rankings to the top 100, and afterwards placing universities into groups of 50 (101-150, 151-200). This means that, however, the rankings themselves lack differentiation between universities not at the very top of the rankings, as the numbers for various metrics will be very small for institutions further down the rankings.

The metrics are geared towards the policy goals of the Chinese government and so as determinants of excellence they reflect a vision that at times appears ill suited to a global view, at other times its metrics appear

somewhat arbitrarily selected. Its subsequent expansion and elaboration can be seen almost as an unexpected side effect of the Chinese modernisation programme. This section will outline the main employment of metrics and recording, and will proceed with a discussion of its strengths and limitations, before looking at USP's performance over the decade that the ranking has been in production.

5.1 Weightings

The ARWU is broken down into four main areas of assessment; quality of education, quality of faculty, research output and per capita performance. For the ARWU FIELD and ARWU SUBJECT rankings the same criteria are used, with slight modifications in weightings and application. Each top ranking university per area receives a score of 100, and each below it receives a weighted score based on its performance compared to the top rated institution.

Criteria	Indicator	Code	Weighting
Quality of education	Alumni of the institution with Nobel or Fields prizes	Alumni	10%
Quality of faculty	Staff holding Nobel or Fields prizes.	Award	20%
	Number of highly cited researchers	HiCi	20%
Research output	Number of papers published in Science or Nature.	N&S	20%
	Number of articles indexed on SCI and SSCI	Pub	20%
Per capita performance	Academic performance per member of faculty.	PCP	10%
			100%

Alumni and Award

There is a general limitation in making cross comparisons of teaching around the world; it is extremely difficult to measure outcomes in different

contexts and different didactic traditions, and the amount of information that would have to be collected makes the exercise nearly impossible. Because of this, the Shanghai team decided to measure education quality in terms of the number of Nobel and Fields prize winners since 1961 the institution counts among its alumni. Nobel winners count only from the scientific fields; literature is excluded from the count. The institution is counted if the winner was awarded a bachelor, masters or doctorate from the institution, and if the faculty member was active at the university at the time the award was made. The institution gains 100% of the credit if the winner studied or worked in the period 2001-2010, 90% for 1991-2000, 80% for 1981-1990 and so on, in this way the ranking makes an effort to reflect current standards and not just historical performance.

First of all, it is not exactly clear how well Nobel prizes accurately reflect the quality of teaching environment in all but a very small selection of institutions. While the United States can count 349 Laureates born within its territory, and the United Kingdom 116, Latin America can count just five that are accepted by the rankings once peace and literature are discounted, and all were granted while the researcher was at a US or British institution. This means that first of all, the metric very heavily favours the established order, and even from that only a small handful of institutions in each country actually hold winners. For those outside of the upper echelons the ranking is not especially sensitive, and especially for those in the periphery who have been traditionally overlooked for a prize that is heavily Eurocentric. This is especially true when it is considered that the literature and peace prizes are excluded, an area in which the developing world is traditionally much better represented than in the sciences.

Because of the scarcity of Nobel and Fields prize winners, the ranking takes a very large time period into account (1961 to the present), and so it is subject to inertia effects. Rather than reflecting current standards, to a certain extent they represent historical ones. This metric does however significantly favour longer established, historically important institutions in the developed world than newer ones in the developing world. The developmental steps made between 1970 (when Nobel Laureates are still taken into account by the

rankings) and the present for the Universidade de São Paulo, which in 1970 was 36 years old and an elitist institution in a military dictatorship, and Harvard, which according to Trow (1973) was already well inserted into a late industrialised and democratic society undergoing massification are very different, and not necessarily a reflection of today's research landscape but of historical importance (Marginson & Pusser, 2013).

Ruminating further on the nature of Nobel and Fields prizes, they are not awarded for consistent high quality performance but for extraordinary, groundbreaking research that changes existing paradigms. As Kuhn (1964) rightly states, such paradigm shifting breakthroughs are by their very nature impossible to anticipate ahead of time. They are, by their definition, moments of genius that disrupt current science rather than fitting in to current paradigms. They represent a very small amount of exceptional scientific output, and rather than defining an entire institution and are better viewed as either personal or team achievements rather than awards that reflect the quality of an institution as a whole.

As a tool for governance this produces a problem. The university cannot formulate policies to produce Nobel Prize winners with any degree of effectiveness, nor are they really markers of research excellence but rather extremely unusual by-products of them. A university can aim for research excellence, from which there may be a chance of a Nobel Prize, and it is certainly true that having a World Class university significantly raises the prospect of winning Nobel Prizes as a result of having top class researchers there, but it cannot introduce a policy to win a Nobel Prize.

Highly Cited Researchers (HiCi)

The ranking attempts to represent the first determinant of Salmi's definition via the number of highly cited researchers in twenty-one specifically chosen subject areas on Thompson Reuters' ISI index 2. Instead of taking total citation rates of researchers, Thompson Reuters focus on only those papers counting as 'highly cited' in the Web of Science Core Collection, defined as within the top 1% of cited papers or articles within the field. This attention to field also is supposed to adjust for the size of the area in which the researcher

operates; the top 1% in some relatively archaic fields like astrophysics will be cited much less frequently than a median paper from a much more mainstream one like clinical medicine, although the relative impact of the former on its field will be much higher than the latter; the HiCi metric is designed to reflect this. By focussing on publications from the last ten years, the HiCi metric is also designed to reflect researchers who are currently influential in their fields as opposed to the net citation rates, which tend to favour senior researchers with a long publishing history behind them, ensuring a more dynamic and up to date representation of influential researchers in their fields.

The HiCi metric avoids the criticisms levelled at the THE and QS for their crude usage of journal impact factor as a proxy for research impact (SEGLEN, 1997), it is not, however, completely without criticism. Various methodological inconsistencies have appeared within the list as a result of misaccreditation owing to identical names of researchers, researchers moving institutions or even passing away (DOCAMPO, 2012). It is also quite obviously and easily 'game-able'; hiring a HiCi researcher included on the list will lead to improvement in ranking position, which in turn has fuelled a 'transfer market' not entirely dissimilar to the football in which teams (or universities) find themselves pulled into an ever spiralling and expensive zero-sum game for a limited talent pool, rather than investing in their own capabilities or having a more spread, or balanced academic staff, it gives incentive to employ a few superstars.

The choice of ISI WoS has been criticised by some for its overwhelming bias towards English language publications as reflected in the Web of Science Core Collection, something acknowledged by its creators who state that because English is the international language of science, they purposefully refuse to index large volumes of material in other languages (PHILLIPSON, 1992; VAN RAAN; VAN LEEUWEN; VISSER, 2011a, 2011b). There have also been suspicions that Web of Science displays favouritism towards Thompson Reuters' own journals, although the second claim would be very difficult to properly verify as a result of the opacity of ISI's own methods. Although the SciELO index is now included in article counts, it is worth reiterating that citation counts from SciELO are not.

Furthermore, the publication of this year's HiCi list revealed an

interesting anomaly that suggests the possible gaming of this metric. It is common practise today for a researcher to list more than one institutional address, as is the nature of modern research; academics are no longer tied to just one institution. This is especially true of highly cited researchers, who tend to be more internationally mobile. However, in 2014's Highly Cited list an extraordinary number of researchers have appeared affiliated with Saudi Arabia as a second address. 82% of Saudi Arabia's 170 highly cited researchers' primary affiliation is with a foreign institution, compared with 2% for the US and 7% for the UK. King Abdulaziz University has a total of 122 secondary-affiliation researchers, with Harvard coming in second place, with 27, while an astonishing 55% of US HiCi researchers with secondary affiliations hold them with Saudi Arabian institutions. USP at present has three entrants on the Highly Cited list, and all three hold secondary affiliations with King Abdulaziz. It appears that by offering financial incentives for de facto academic positions Saudi Arabian institutions have artificially inflated their performance in this list by buying the prestige of other researchers without producing any of the attendant research (GINGRAS, 2014). Given that academics tend to forego much greater financial rewards in the private sector given their expertise, it seems only natural that when an institution approaches with a financial bonus to do ostensibly nothing, academics will accept. This should not be taken as a denunciation of those who do accept the money, rather as an example of the fallibility of the metric.

This seems to correlate rather closely with King Abdulaziz's stratospheric rise through the rankings from the 400-500 group in 2012 to the 100-150 group in 2014, moving from a score of 0.0 as recently as 2012 in HiCi to a jump to 31.6 for 2013 and 33.1 in 2014. The slow growth and small variation for other institutions, and indeed in its performance in other metrics, which have not grown at anything like the same speed would certainly suggest at the very least unnatural growth, and probably the presence of gaming behaviour. This shows that depending too heavily on quantitative metrics to measure quality gives incentives for corrupt practise (CAMPBELL, 1976).

Nature and Science (N&S)

This section, designed to measure excellence and impact of research, is made up of the number of articles published in the journals *Nature* and *Science* as a proxy for high impact research that reaches beyond disciplinary boundaries to the wider scientific community. While it is undoubtedly true that higher rated research tends to be published in high-ranking journals, it is not entirely clear why *Nature* and *Science* were chosen specifically to represent this. While both are extremely well cited in general, and usually represent the very top of the field when published (see the co-citation network produced in the final section of this dissertation), it seems arbitrary and exclusive to discount the entirety of the rest of the publishing landscape for these two. Why, for instance, PloS One was not selected too is impossible to justify.

This metric is susceptible to the criticisms levelled at use of journal impact factor ranking as a proxy for quality (BREMBS; BUTTON; MUNAFÒ, 2013; TOGIA; TSIGILIS, 2006), arguably to an even greater degree. By placing so much weighting on just two journals, albeit well respected multidisciplinary ones, leads to massive dependence on a relatively arbitrary indicator, one which represents a miniscule proportion of total published work.

Publications Indexed on SCIE and SSCI (Pub)

The Pub metric is designed to measure academic output in the Science Citation Index- Expanded (SCIE) and Social Science Citation Index (SSCI), both produced by Thompson Reuters. It takes into account the total number of papers indexed for the year of the ranking. This metric intends to serve as a representation for the year's scientific output, rather than as a total historical representation of the university's published output. In this way the ranking attempts to balance some of historical and geographical criticisms levelled at it with regards to the Alumni and Award categories, and of the problems of comparing across disciplines covered in this dissertation.

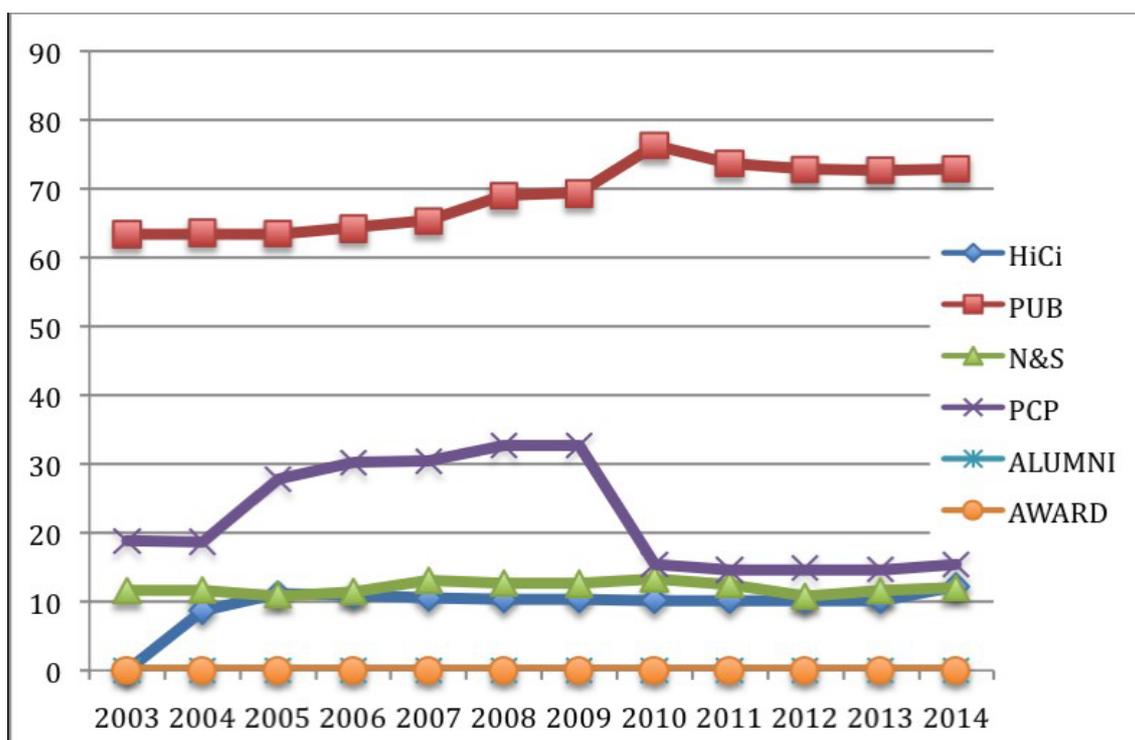
Per Capita Performance (PCP)

This metric is a combination of the previous four divided by the number

of full time academic staff the university employs. The aim here is to show how productive the university is. This metric will clearly favour small, research-intensive universities like Caltech over larger comprehensive universities more commonly found in Latin America. In the higher scoring institutions on this metric the whole academic staff are directly engaged with the research process, generally fitting more American models of university than the 'nation building' Latin American ones with heavier pedagogical commitments than their northern counterparts.

5.2 USP's Performance in ARWU

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
HiCi	0	8.7	11.1	10.9	10.5	10.3	10.3	10.2	10.2	10.2	10.2	12.2
PUB	63.4	63.6	63.5	64.2	65.3	69	69.3	76.3	73.7	72.8	72.7	72.9
N&S	11.6	11.7	11	11.5	13.1	12.7	12.6	13.3	12.5	10.8	11.6	12
PCP	18.8	18.7	27.8	30.2	30.4	32.7	32.7	15.4	14.6	14.9	14.7	15.4
ALUMNI	0	0	0	0	0	0	0	0	0	0	0	0
AWARD	0	0	0	0	0	0	0	0	0	0	0	0



As a result of the aforementioned grouping of universities outside the top 100, USP has retained a relatively stable position in the overall rankings, moving from the group 151-200 to 101-150 in 2004, and remaining there ever since. This is perhaps as a result of the extremely small number of new Nobel Laureates each year, meaning that at least 30% of the weighting is highly unlikely to move for all but a few institutions each year. Because of its position outside the top 100 the aggregate score is not reported. It is worth noting here that as these are the cardinal weighted z-scores, and not the ordinal ranking position scores, a higher score denotes a better performance.

As might be expected from Chaimovich's analysis (2014), Brazil's 'adolescent' phase of research, and its associated focus on volume of research in both CAPES evaluations and the Anuário, has produced remarkable results in the volume of research produced in the Pub metric. At launch in 2003 the institution already registered an impressive 63.4 as a score weighted against the top ranking institution, placing it higher than Yale (61.1) or Princeton (62.5). As a result of the privileging research output over impact in evaluation, by 2013's ranking USP recorded a score of 72.7, placing it fifth overall in the rankings. This is a tremendous achievement in line with the university's institutional goals and broader national goal to increase scientific output and serves to demonstrate the university's considerable productive capacity.

A different picture is painted by USP's performance in the HiCi metric of highly cited researchers. An initial score of 0.0 in 2003 rose over the following decade to 10.2 in 2013. However, as a global score, this still leaves the university in 287th worldwide, comparatively lagging behind most of the rest of the 101-150 group in terms of research impact. In the 2014 Highly Cited list, USP counted just two researchers within its ranks, both from the department of chemistry. Similarly in the other metric designed to measure research impact, N&S, a score of 11.6 places the university in 301st position globally, again relatively much lower than the university's place on the ranking, and lower than many of the universities in the same group. While these are not universal or objective measures of impact, they should be of concern to USP, if these particular partial indicators show such poor performance, it reflects much current discussion of USP and its need to increase the visibility and impact of its

research.

5.3 Conclusions from Shanghai Jiao Tong ARWU

What the Shanghai Jiao Tong shows us about USP is a vision of an institution with enormous productive capabilities, that compares extremely favourably with World Class universities in terms of output, but one that struggles to translate this output into academic bibliometric impact. While there are a huge number of indexed articles on SCIE and SSCI from the Universidade de São Paulo, there are very few reaching highly cited status. This is in line with many commentators' views that Brazilian higher education needs urgently to rethink its approach to research orientation and impact in line with its societal obligations to produce research that has an impact upon the both society in which it is embedded and the broader global community.

It can be countered on the other hand, that the Shanghai Jiao Tong measures impact of research in a constrained and perhaps unrepresentative way. First of all it considers purely bibliometric data; it measures only academic as opposed to social or economic impact. Furthermore, in focussing on only the top 1% of researchers in any given field, focuses more on individual achievement than all round departmental strength, and lacks granularity to measure universities outside of the top twenty. It also privileges models that conflict profoundly with the Latin American 'flagship' model by favouring research intensity spread across the whole institution. This is unsurprising as the original remit of the ranking was to determine the world's best research-intensive universities as defined along the American model, rather than provide a detailed vision of world academia. Furthermore, its measuring of teaching and faculty excellence is restricted to Nobel Laureates among its staff and alumni, something that lends a heavy bias to historically established institutions, and to institutions in the developed world, and therefore are not necessarily a good indicator of all round quality, but of individual excellence.

6. TIMES HIGHER EDUCATION AND QUACQUARELLI SYMONDS RANKINGS

As a result of the striking similarities in structure and intention between the two, and their difference from the ARWU, the Times Higher and Quacquarelli Symonds will be broadly treated together, with the significant differences between them highlighted.

In 2004, the year after the first Shanghai Jiao Tong ranking, the British newspaper Times Higher Education (BATY, 2010) published its first global university ranking as an extension of its national ranking for British universities in conjunction with the educational management consultancy firm Quacquarelli Symonds, responding to the fact that higher education systems are no longer measured purely nationally, and increasingly elite institutions wish to compare themselves internationally as World Class universities rather than merely successful domestic ones. It is what is known as a composite ranking; rather than measuring external bibliometric research information like the Shanghai Jiao Tong it aims to give a more rounded view of a university's quality by taking teaching, research and reputation into account. The THE national rankings are geared largely towards consumer information for undergraduates, intended as a buyers' guide to higher education, rather like the domestic RUF ranking. This commercial information focus underpins both the THE World University Rankings and QS as well, however the focus has moved from domestic undergraduates to a more global outlook, with the expressed intention of attempting to cross information gaps between higher education systems in a more globalised world (Baty, 2013).

Indeed, this consumer focus is reflected in the fact that Quacquarelli Symonds holds its annual global MBA tour, organising fairs for prospective students and directly positioning itself in an advisory role to the young, high-achieving and privileged cosmocracy of discerning and internationally mobile students. In this way, its rationale is fundamentally different to that of the ARWU, and this is reflected in its make up of metrics. It does not, like the ARWU, aim to give an account only of the top of the rankings, but allow for

cross comparison of a variety of institutions and give a better description of international higher education.

The THE underwent a fundamental methodological change in 2009 in ending its six year relationship with the consultancy firm Quaquarelli Symonds, who had supplied and interpreted the data from Thomson ISI and administered the reputational survey. The split came following the some of the methodological concerns highlighted below, where the THE has attempted, very publicly, to right some of the inconsistencies still present in the QS.

Reputational Surveys and their Limitations

Both the QS (O'LEARY; INCE; AL HILI, 2010) and, to a slightly lesser extent the THE rely on a reputational survey of academic and research quality. This can be seen as an attempt to introduce a qualitative judgment element into the predominantly quantitative dimensions of bibliometric rankings. This aligns itself better with the observation that there is no one indicator of quality; rather it is a multifaceted phenomenon with an unstable definition. In this respect, it is seems desirable to allow professionals make their own qualitative judgments on elite institutions and aggregate their responses in this way.

However, reputational rankings have become the most contentious of all metrics in academic circles, for two broadly related areas of complaint; the normative/power dimensions implicit in the measurement of status relations, and the not inconsiderable methodological limitations of such an assessment. Altbach goes as far as to observe that the more weighting given to reputational surveys, the less we may depend upon the ranking as a policymaking tool (BOWMAN; BASTEDO, 2011; MARTINS, 2005; SAFÓN, 2013; SALMI; ALTBACH, 2011).

The rise of bibliometrics as a method of evaluation closely corresponds to a need to overcome the cronyism that often blights qualitative and departmental or school level evaluation. It gives a way that public observers armed with a little specialist knowledge, may draw conclusions and make judgments about the productivity and impact of research.

Reputational rankings tend to favour a certain profile of institution.

Universities with strong overseas recruitment and delivery profiles usually fare better than institutions that focus more on domestic delivery. This is why the rankings appear to favour Australian universities more heavily than Canadian ones, despite the fact that in terms of systemic health and academic profile Canadian institutions are arguably slightly stronger.

It is especially telling that, in response to Gladwell's *New Yorker* editorial on the relative uselessness of reputational rankings (GLADWELL, 2011), the head of Times Higher's ranking department explains how it has attempted to overcome some of the methodological weaknesses, and that the Times Higher now depends less on reputational surveys than before, without addressing these underlying normative issues related to their use.

In the interview with prof. Nunes he revealed that one of the few incentives that USP takes as a direct response to rankings is to heavily invest in international academic visits and guest lectures as a way of boosting the university's international presence. While this increases the amount of interface that the university has with the outside world, it does also point to the role that cronyism has in this kind of measurement. The amount of time a visiting lecturer actually has to assess a university in one, perhaps two days of visits is limited, however when presented with a voting slip, they are much more likely to select the university that acknowledged their research and with whom they have some personal connection.

Serious questions have been raised about the methodological soundness of reputational surveys. First of all is the problem of respondent reliability. The extent of an academic's knowledge about the intimate inner workings of other institutions is questionable. Although academics are asked only to comment on institutions of which they have intimate knowledge, the reality is likely to be that knowledge of institutions is likely to be uneven, and therefore unsuited to an ordinal ranking system. If academics are only permitted to evaluate institutions with which or in which they have worked, this is likely, for the vast majority of any sample, to exclude many institutions generally regarded as quality. What happens in practise then, whether or not an academic knows specific information about an institution, they evaluate according to generally accepted status.

When QS began its subject rankings in 2009, for example, Harvard gained third place in the rankings for geography, which is an impressive achievement for an institution that famously closed its geography department in 1948 with the statement that 'geography is not a serious subject of study' (SMITH, 1987). Although geography covers a large interdisciplinary space, much of which does cover what Harvard does, it is doubtful that this crossover can take the place of a fully operational department. In the English Literature section of the same year, the top 20 included Peking University and the Chinese University of Hong Kong, which reflects the tendency pointed out by others to artificially promote regional Asian leaders in reputational rankings as a result of the heavy skew in statistics caused by the predominance of Asian respondents in reputational surveys, where rankings and prestige are much more central concerns to policymaking than in the West.

This is not to suggest that Peking or CUHK are not excellent universities in their own right, but having universities where English is not the first language nor the language of course delivery rated as stronger than many traditionally strong Anglophone departments seems at best surprising and counterintuitive.

To compound this matter, amid widespread ridicule at the first publication, none of these anomalies appeared in any of the subsequent QS subject rankings. This suggests that either the academic community had a collective damascene moment of clarity and altered their voting habits on the basis of becoming better informed in the space of a year, or that the results were expunged from the count in the face of QS's public embarrassment, adding fuel to the claim that commercial rankings' main aim is to reconfirm status quo as a way of maintaining legitimacy rather than being an accurate measurement of performance (HUANG, 2012). When Altbach's centre-periphery conception is taken into account, the centre characterised by dense networks of actors, with a relationship to the periphery, and the periphery defined by much looser associations with one another, but stronger links with the centre we would expect that the outcome of reputational surveys in this case to promote Anglo-American institutions heavily active in the region, and a few regional leaders. They would typically not take into account other

'peripheral' institutions, such as non-Anglophone Europe, or Latin America. This is exactly what we see in the results of both the QS and the THE reputational rankings.

Furthermore, as a result of the volatility present in reputational rankings, it has been suggested that they present too much statistical noise to be of much use in policymaking, particularly as the methodology is tweaked or altered every year, and especially for the THE in 2009-2010, when it was changed altogether (BOOKSTEIN et al., 2010). This makes year-on-year comparison difficult to judge, as these methodological changes are presented somewhat ambiguously.

6.1 Times Higher Ranking

The new THE uses a set of thirteen different metrics in an attempt to represent a broader variety of institutional aspects. They can be broken into five separate areas; research volume, citation and impact, teaching, internationalisation and industry income. In an attempt to take into account the diversity of institutional profiles in global higher education, the THE produces Z scores based on deviation from the mean score of the dataset before producing a cumulative probability score for the metric, allowing for theoretically fair comparison of data. However, while it can statistically control for size, there are still various benefits associated to institutional models and sizes that cannot be statistically adjusted. For example, smaller institutions like Caltech have an added benefit of being extremely small and with limited responsibilities towards teaching or extension. This means that even when normalized for size, the range of functions performed by USP academic staff across different units is far broader than in somewhere like California Tech or IMPA in Rio de Janeiro, which can maintain an extremely intense focus on research.

6.1.1 Teaching

Due to the problems associated with cross cultural comparison of teaching styles, outcomes and employability, the ARWU continues to disregard teaching measures in its ranking. The THE however, as it is primarily aimed at prospective students, has attempted to represent the teaching dimension.

Teaching output is by its nature context dependent; employability, for example, is dependent on the job market in which the university finds itself, and is therefore not an objective measure of teaching quality. In response to this the THE chooses to measure the teaching ‘environment’, in an attempt to construct proxies for quality, or at least give the user an impression of what could be expected from the learning environment. In total it is worth 30% and is broken down into the following;

Metric	Methodology	Weighting
Teaching Reputation	Global Survey	15%
Postgraduate academic capacity	PhDs granted per academic	6%
Undergraduate academic capacity	Undergraduates admitted per academic	4.5%
Financial Resources	Total institutional income/ Total academic staff	2.25%
Research intensiveness	Ratio of PhDs awards to Bachelors	2.25%
Total		30%

Half of the teaching metric (15%) and roughly equivalent in the research metric (18%) is given over to a reputational survey, as we saw with the QS ranking, although according to Baty (2013), it has been restructured for appropriateness with a series of structured questions to get more qualitative details from respondents. However, the same criticisms that arise in the QS still hold here; there is no transparency in the survey, no transparency in the respondents’ details, little control over sampling over geographical and disciplinary boundaries, no accountability and no reproducibility. The new THE has gone some way to addressing the criticisms levelled at the old QS methodology; that it excessively favours Northern institutions and established ‘elite’ institutions and therefore is simply a reflection of halo effects rather than being a record of genuine academic excellence. According to the Global Perspective on Research and Development report (UNESCO, 2010) the distribution of the world’s academic population is approximately;

North America	22.20%
Europe	28.40%
Asia	41.40%
Oceania	2.10%
Latin America	3.60%
Africa	2.30%

While the distributional spread of the THE sample for 2013 was;

North America	33%
Europe	26%
Asia	23%
Oceania	10%
Latin America	5%
Africa	4%

For Latin America and Europe this is a reasonably representative sample, however there is a huge imbalance between North America and Oceania on the one hand, and Asia on the other, which is likely to produce a skew heavily in favour of English speaking countries, and even more specifically towards the US elite institutions. This is further compounded by the fact that North American academia tends to be characterised by much denser collaborative networks than other regions, making it proportionally more likely for respondents from North America to have more experience of their immediate neighbours, and therefore make them more likely to indicate them over departments in other regions.

The survey is more finely attuned to disciplinary specifications than the QS, asking academics to indicate specifically strong departments and not institutions with broadly strong areas to prevent elite institutions coming out top by default. However, there is still a perhaps unintended distortive effect across disciplinary boundaries in that response rates from social sciences and STEM subjects were much higher than from humanities, perhaps stemming from a lack of confidence in ranking exercises in these areas. This means that the rankings will typically favour institutions with scientific focuses over those with

broader or more social focuses.

Furthermore, the THE, through Thompson Reuters, subcontracts a third party it describes on its website as a professional polling institute to conduct the survey. Upon closer inspection this third party is Ipsos MediaCT, not to be confused with the respected British pollsters Ipsos Mori, although they are part of the same organisation. Ipsos MediaCT is not a polling company, but an online digital marketing agency specialised in analysing media usage, brand management and consumer audience profiling. This gives serious doubts as to the methodological rigour and oversight either the THE or Thompson Reuters has over its reputational survey, as market research companies typically present final results without allowing a lot of involvement with the data collection process. They are not held to the same standards of investigation as either pollsters or academic surveys, and so their presentation as academic surveys is erroneous.

The teaching environment is also assessed the ratio of faculty members to students, with the assumption that the fewer students per member of staff leads to a more attentive learning environment for students, and therefore better learning outcomes. As it is divided in the THE between admitted undergraduates per academic and PhDs granted per academic, this is supposed to represent a higher degree of granularity in how research intensive the teaching environment is. However it seems peculiar that the THE has chosen to measure one input (undergraduates) with one output (PhDs granted) next to one another. One possible explanation would be the risk of overfilling at entry in the expectation of a high dropout rate at undergraduate, as USP suffers from. This metric has drawn much criticism as a measure of quality of environment, because it is very difficult to verify and very easy to 'game'. The definition of academic staff is never fully clarified as a result of the diversity of academic titles and grades present across cultural contexts, as well as different employment practises.

There is, for example, a huge difference between a fully tenured senior professor delivering lectures and a doctoral candidate overseeing undergraduate seminars. It is not entirely clear whether or not postgraduate students on USP's PAE programme or part time lecturers are counted as

academic staff in the count, as the company does not explicitly use any of the international human resources in science and technology (HRST) scales, such as the ISCED or ISCO (UNESCO, 2011; ILO, 2007). It has become increasingly prevalent practise for many universities worldwide to use doctoral and post-doctoral students instead of academic staff for many academic duties, meaning that the proxy becomes very vague as a measure for attentiveness. This in turn promotes the adoption of perverse incentives for policy makers, as the use of three teaching assistants for the money that could be used for one lecturer means a much better performance on the ranking using less qualified staff. This in turn is having a destabilising effect on the academic trajectory, making the step into an academic career more difficult.

To some extent the THE recognises this problem, and admits that it can only serve as a very crude proxy for teaching environment quality (BATY, 2013). In recent years it has significantly reduced the weighting of the staff-to-students metric; it now represents just 4.5% of the total, meaning that any gains from reengineering employment policy in order to game the metrics would be very marginal.

The ranking measures the research intensiveness of an institution by the number of doctoral students in comparison to undergraduates, based on the observation that a greater mix of postgraduate students gives a more research-intensive learning environment to both undergraduates and postgraduates. It also implies a commitment to training the next generation of researchers and academic community rather than simply being post-secondary learning environments. This focus on the postgraduate environment is further enhanced by measurement of the total number of doctoral degrees awarded in the university compared to total number of academic staff. This number is produced as a ratio in order to give a representation of the focus an institution gives to postgraduate education rather than as an absolute figure, which would be to a certain extent a measure of institution size and not environment. Both of these metrics, rather like the PCP measure of the ARWU, heavily favours the Anglo-American model of Research University in which research focus is spread evenly across the academic staff, ignoring other models of higher education in which research is concentrated to some sectors, while others fulfil more

pedagogical roles.

6.1.2 Research Volume, Income and Reputation

The research volume metric is subdivided in a similar way to the teaching metric, and is worth 30% of the total:

Metric	Indicator	Weighting
Research Reputation	Opinion survey	18%
Research Intensity	ISI-Indexed articles/academic staff	6%
Research Income	Total grants awarded scaled by discipline and size	6%
		30%

Research volume is collected through data indexed on Thompson Reuters' ISI index. The THE scale theirs as a ratio compared to total number of staff, instead of the net figure. In contrast with the ARWU, the THE does not measure total research output, but is more interested in research density, favouring universities that spread research throughout the entire institution rather than concentrating in some areas.

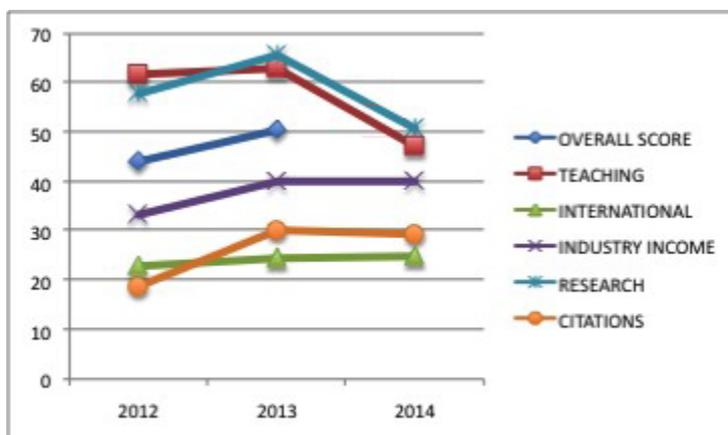
At 18%, research reputation is the second largest single metric after citation influence and impact. It is unclear however, what dimension a reputation for producing high quality research represents that cannot be represented by bibliometrics. Admittedly, not all citations are good citations, as often a paper will be cited criticising it, and certainly not all highly cited papers are good papers. However, the correlation between producing high quality, high profile research and having it highly cited is, to say the least, high. High citation rates demonstrate that for better or worse the research has had significant impact on global discussions about research. The reputation of research however, is likely to favour exactly the historical continuity that the Shanghai Jiao Tong is trying to avoid with its metrics. It is not necessarily true that whoever an academic has cited most often or read most of during a year is what they necessarily regard as the most esteemed, indeed it is perfectly possible that they have read very little in years but are in fact simply reflecting commonly held opinion. Like the QS, there is a strong suspicion that what is in fact being represented are halo

effects attached to large and reputable institutions rather than being a measure of the actual research being carried, or an institution’s current role in world academia.

6.1.3 Data collection system

In order to facilitate the transfer of large amounts of data from universities to Thomson Reuters in compiling these rankings they have developed a digital platform in order to build more comprehensive and accurate institutional profiles, allowing for better comparison and the prospect of benchmarking institutions in the future rather than ranking them (SALMI, 2013). The next chapter on the university’s Anuário Estatístico will look in detail at whether or not its data collection methods and processes are compatible with external measurement exercises. For years of non-reporting of information the ranking is supposed to give a simple mean score based on the information from previous years and peers within the same category to attempt avoid punishing institutions for non-compliance. **USP’s Performance in THE**

	2012	2013	2014
OVERALL SCORE	44.1	50.5	Withheld
TEACHING	61.7	63	47.2
INTERNATIONAL	22.9	24.5	24.9
INDUSTRY INCOME	33.3	40	40.1
RESEARCH	58	65.7	50.8
CITATIONS	18.8	30.2	29.4



As the Times Higher is still relatively young, long term comparisons are

difficult, but it is clear that the scores for are much more volatile than in the ARWU, with large variations in scores between years, meaning that its utility as a policymaking tool is somewhat limited (BOOKSTEIN et al., 2010). Furthermore, as the make up of each metric is itself compound (it has several constituent components), meaning that the actual cause of the fluctuations is not necessarily clear. For example, the huge fall in 2014 of research and teaching was provoked by any one, or combination of, three and five metrics respectively. As before, these are the cardinal not ordinal scores and so a higher value denotes a better performance.

This fall was probably not, however, significantly precipitated by a fall in the reputational rankings, as the Times Higher has taken to publishing them separately as the Global Reputational Ranking. USP performs disproportionately well in these, placing consistently within the world top 100, although for this year there was a small decline, moving from the group 61-70 to 71-80. This would suggest then, that the fall was caused by one or more out of the research income, research intensity, financial capacity, or one of the postgraduate intensity metrics.

6.2 Quacquarelli Symonds

The Quacquarelli Symonds, as previously mentioned, is run by the eponymous British consulting firm, whose concerns are more related to the commercial, consumer side of rankings than perhaps the other two are. As such, their metrics are much less focussed on bibliometric performance indicators and much more on reputation and institutional perception.

Metrics

Metric	Indicator	Weighting
Academic reputation	Global reputational survey of academics	40%
Employer reputation	Global reputational survey of employers	10%
Teaching environment	Student to faculty staff ratio	20%
Academic impact	Mean of citations per faculty	20%
Internationalisation of staff	Ratio of international academic staff to domestic	5%

Internationalisation of student body	Ratio of international students to domestic	5%
--------------------------------------	---	----

The same caveats previously discussed on academic reputational surveys apply for the QS, with the added note that a much higher weighting is given over to it in the QS than in others. This clearly favours universities with international marketing and profile, and those who pursue activities such as transnational course delivery, aggressive foreign recruitment and foreign campuses, predominantly the domain of Anglo-American institutions.

6.2.1 Employability

Employability is difficult to measure across national contexts, as the employment rates or salaries of alumni are so intrinsically tied to the national or regional economic situation of the state in which the institution resides. As a result, these types of measurement are not valid performance indicators for global rankings, as they might be for national ones. Similarly, due to the absence of internationally recognised standards for evaluation of students entering tertiary education, 'value-added' scores that aim to control for selectivity by not excessively favouring institutions who are highly selective have limited applicability on international rankings. In recognition of this, the QS places all of its weighting on an employer reputation survey.

It has been suggested that employers very often lack the specialist knowledge to make quality judgments about the often complex inner workings of an institution, and therefore rely more on outside perception and historical weight than any intimate knowledge of quality. This means that they are more likely than academics to be led by halo effects from historically strong institutions, and more likely to refer to previous years' rankings for guidance rather than making actual value judgments. While this may be a helpful indicator for prospective undergraduate students, its utility is limited for institutions themselves, and not a measure of quality per se, but of perception.

6.2.2 Teaching intensity

To assess teaching environment, the QS relies solely on the faculty staff to student enrolment ratio. This metric is prone to gaming by administrators because it measures input rather than output or impact. The numbers can be adjusted with the addition of teaching assistants and post-doctoral researchers given the same weighting as fully tenured professors. It is also predicated on the assumption that all, or at least the majority, of students will study on a full time basis. At USP, the presence of multiple sessions of students, during the day and night, and a large number of part-time and extension courses means that while the absolute student number is vastly increased, it does not necessarily signify much larger classrooms. It also does not necessarily follow that more academics will signify better teaching, and as such will not measure quality of teaching environment. The metric could be improved in its accuracy by measuring a ratio of the hours the academic staff is contracted for to hours of student contact time. This would help to control for the diversity of course and system, and reduce some of the gaming incentives. It still would not measure teaching quality, but it would better represent 'intensity' of the study environment.

6.2.3 Impact

QS structures its impact metric with the vague description of citations per faculty using Elsevier Scopus instead of Thomson ISI, claiming that Scopus can boast better coverage of international journals, although this claim is somewhat controversial, with most research being fairly inconclusive (HICKS, 1999). In the course of collecting data for the final chapter, it was found that Scopus returned a much higher number of results for USP, and so for much of the analysis it was preferred, but this is purely anecdotal and should not be taken as any more than that.

6.2.4 Internationalisation

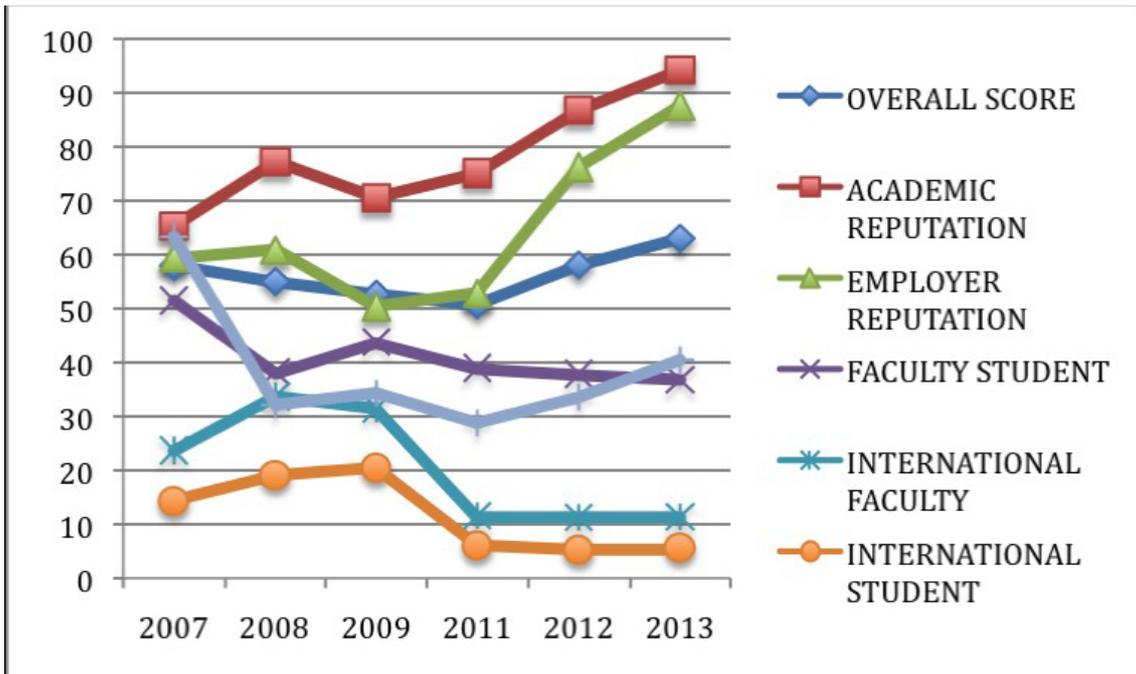
Because a World Class university has the ability to attract and retain top global talent and has recruitment capacity beyond its borders, the QS also

evaluates internationalisation as a proportion of both teaching staff and student body. Measuring this diversification of the student and staff body is a positive element of this ranking, as it stimulates more mobility of staff. However the literature has been unable to establish a causal link between mobility and performance; whether mobility stimulates better performance, or whether better performance causes academics and students to be more mobile (CAÑIBANO, OTAMENDI, SOLÍS; 2011).

In the publication of this year's metrics, the QS is much more sanguine about the limitations of its metrics than in previous years, clearly stating some of the limitations of what is measured and why it is measured. This is possibly as a result of considerations brought about by the IREG accreditation.

6.2.5 USP performance in QS

	2007	2008	2009	2011	2012	2013
OVERALL SCORE	57.9	55	52.66	50.7	58.06	63
ACADEMIC REPUTATION	65.3	77.2	70.5	75.1	86.8	94.2
EMPLOYER REPUTATION	59.4	60.8	50.4	52.8	76.2	87.6
FACULTY STUDENT INTERNATIONAL	51.5	38.3	43.8	39	37.9	36.8
FACULTY	23.7	33.7	31.5	11.6	11.3	11.4
INTERNATIONAL STUDENT	14.3	19.1	20.5	6.1	5.3	5.7
CITATIONS PER FACULTY	63.3	32.1	34.3	28.8	33.6	40.5



6.3 Combined Performance Analysis

If rankings are as objective and externalized as they claim to be, we should be able to expect that movements in one will be reflected in the other two, and borne out by evidence from the Anuario Estadístico. Because all three of the rankings and the Anuario, measure information in different ways this is not as simple as merely lining them up on a graph to look for correlation. However, what we would want to see is whether or not an increase in scientific production appears in just one, or across the board, whether significant increases in reputation scores is the result of better performance in the previous year, or some other factor unrelated to performance.

6.3.1 Scientific output

In the ARWU, which measures the number of articles on the SSCI and SCIE indexes, USP's performance over the past 12 years is the following:

PUB	63.4	63.6	63.5	64.2	65.3	69	69.3	76.3	73.7	72.8	72.7	72.9
-----	------	------	------	------	------	----	------	------	------	------	------	------

This is a steady, but significant increase in z-scores relative to the top ranked institution. This would suggest that scientific production was increasing in the university, but in the Anuario Estadístico, we can see that scientific production across all types of publication is actually declining:

2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	
29105	28505	28724	30992	29882	30097	29347	28425	27211	25425	25653	

To explain this seeming anomaly, we can point to a number of possible explanations. Perhaps there is a global trend towards publishing less, but of higher quality, meaning that USP is behind the trend compared to top institutions. Publishing behavior could be changing within the university, with academics more inclined to attempt to publish in indexed journals than other sources. This is impossible to tell from the anuario as it does not measure these statistics. Most likely, as the SCIE and SSCI indexes have extended significantly in recent years, they now include many more local journals that were not previously indexed. If it is this, then the progress this metric seems to represent is in fact a measurement not of the university, but of Thomson Reuters' measuring ability.

Likewise, in productivity measures (PCP in ARWU) we see from 2003 to 2014:

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
PCP	18.8	18.7	27.8	30.2	30.4	32.7	32.7	15.4	14.6	14.9	14.7	15.4

While according to the anuário, production per capita reads:

	2009	2010	2011	2012	2013
Média de trabalhos por capita docente	5.6	5.6	5.3	4.8	4.8

While these are relatively stable in both, we are at a loss to explain how, when according to the anuário the mean number of papers per faculty member

remained exactly the same (5.6), the score in ARWU decreased by more than half. This strongly suggests that this dramatic change is not precipitated by internal factors, but more likely some methodological change, or external factor not taken into account here.

6.3.2 Reputational Surveys

For reputational surveys to be valid, especially when comparing year to year, we would expect large jumps in score to follow a year after large deviations in performance metrics, and for large changes in one to be mirrored by large changes in the other. This is signally not the case when looking at the QS and THE reputational rankings.

For QS, from 2007 to the present, the academic survey results have been

ACADEMIC REPUTATION	65.3	77.2	70.5	75.1	86.8	94.2
---------------------	------	------	------	------	------	------

While for THE over the past three years, the academic compound scores, of which 18 of the 30% is made up of the reputation survey scores, read:

RESEARCH	58	65.7	50.8
----------	----	------	------

Whilst we know that scientific production did not decline or rise significantly according to the ARWU, and citations per faculty have apparently increased according to the QS ranking (final three columns);

CITATIONS PER FACULTY	63.3	32.1	34.3	28.8	33.6	40.5
-----------------------	------	------	------	------	------	------

This means that the only feasible explanation the wild deviance in the compound score in the THE is the academic survey, which in the reputational ranking is withheld, we can see that the university dropped several groups in 2014 before returning to the same level in the 2015 ranking. From this, not only

do we not see significant correlation between rankings performance, but we do not see visible external reasons as to why reputational rankings rise and fall so wildly. While the survey showed fantastic results in the QS in 2014 (94.2), it showed terrible results in the THE (50.8 for all academia). This ought, at the very least, lead us to question the validity of conclusions that can be drawn from these rankings.

7. THE UNIVERSIDADE DE SÃO PAULO

The main hypothesis of this project is that the Universidade de São Paulo is not able to take full advantage of new mechanisms for international evaluation as a result of its governance structure and the way in which it records information about itself. As we have established in the previous section, international evaluation offers some important opportunities for the future and modernisation of Latin American higher education, as well as some profoundly troubling tendencies. The challenge for policy makers in so-called flagship universities is to harness this potential in a way that enhances quality in line with institutional and regional aims without compromising themselves in pursuit of goals dictated by external ranking agents.

7.1 Anuário Estatístico

This project of international insertion and evaluation, essential for an institution seeking to be a regional leader or world class university, essentially rests on the ability of its executive to record and monitor activity within the university in order to respond to the realities of the laboratory or lecture theatre and orientate policy in a way that reflects this. The Anuário Estatístico attached to the reitoria performs this role at the Universidade de São Paulo (UNIVERSIDADE DE SÃO PAULO, 2013).

The Anuário measures an array of information about the university, from its technology, library systems, physical upkeep, budgets and expenditure, cultural extension, social engagement and responsibility to teaching, CAPES evaluation and research metrics. In this way the Anuário aims to first of all serve as a support tool for management and decision-making, but also in a wider sense to represent the 'life' of the university. As a tool for increasing public accountability of the university it is indispensable if it represents an accurate vision of the university's activities.

This analysis should be taken with the important caveat that the Anuário was a project undertaken under the auspices of prof. Goldemberg's administration to publicly represent something of the university's activity. It was

undertaken on a voluntary basis by members of university staff based on the statistics they had to hand, and it is in this vein it continues to this day. It is also very unusual in university administration to be so transparent with information. Any criticism of the Anuário must be considered in this light, not that it is inadequate, but that it could be something much more extensive than it is.

Care must be taken not to over analyse the importance of what is, and what is not contained within it, as it is not necessarily used at the centre of the university's strategic planning committees. What this section aims to do instead, is to demonstrate how the Anuário could be better professionalised and used in a more integral way in policymaking or for public communication.

As this project is specifically about rankings and international insertion, only the years 2003-2013 will be considered, although the Anuário was established in 1987 with the granting of full institutional autonomy. Furthermore, only certain parts of the Anuário will be considered here as a result of the relatively limited measuring focus of international rankings today. However, the areas chosen are specifically those that USP stands to gain most from greater engagement with international evaluation; section three (Academic Information), section five (library system), section ten (International Relations), and to some extent sections two (Demographics) and four (Publications).

7.1.1 Section Three: Academic

Tables 3.27-3.33 form the base of how the university measures academic production, and therefore the productivity of its units. However, the numbers reported varies hugely for the same year in successive Anuários, suggesting that the collection of publication data in some cases takes more than the allotted year, or to changes in collection methodology. This complicates the process of comparison significantly, as there are several figures available for each year, leading to possible confusion. This suggests that perhaps the annual census format is not the best for measuring scientific production in this way. In 1987 this was the only feasible option, but today perhaps a more dynamic continuous system based on self-archiving would be more appropriate, given the availability of the technology to achieve this.

As a result, the figures taken are of the most recently recorded; the

2003 statistics are taken from the 2008 edition, but 2008 is taken from the 2013 edition, whilst the years 2009-2013 are all taken from the 2014 edition. This means that there is a chance that some of the figures from more recent years will appear lower as a result of delays and problems with reporting.

The first table of analysis is a breakdown of academic production by type and location of publication. The usage of type gives a much broader spectrum of academic activity than considered within rankings, which because they use either ISI WoS or Elsevier Scopus, consider indexed, refereed journals only. As discussed earlier this leads to distortive effects in judgment of research as it ignores the majority of published conference proceedings, books and a variety of other publishing media. It also gives a breakdown of publishing by location, national or international. While neither of these sections refer to impact or reach of research, they provide a useful account of how the publishing habits of the university as a whole's publishing habits are changing over time, something which is not necessarily clear from looking purely at rankings. For instance, we can see a general decline in the total volume of published output over the last decade, but with a corresponding increase in the output published in journals, possibly as a result of the rise of the open access journal, and a definite move towards more international publishing. This suggests that behaviour is moving towards a more impact-orientated approach, but at present the measurements used are not in a position to qualify this claim.

The next table in scientific production is by department in terms of book and article publishing, and then in conference proceedings and annals respectively. While this is useful to track some of the various different publishing behaviours between departments, lacks the granularity to fully capture publishing behaviour. The category books and journal articles is too broad for the purposes of rankings and impact measurement, many disciplines favour book publishing over article, and while this may give some semblance of an overall view of output, it does not do it in a way that is necessarily constructive.

Something that the anuário ought to take into account in this section is the total value of public funding received for research for each unit. This would include total resources captured from FAPESP, CAPES, CNPq and foreign public investment. It would greatly help transparency in income, but more

importantly highlight areas within the university successful in attracting public funding for its activities, to allow exchange of information, shared learning of successful policies and areas active in publicly funded work. We could also, if some effort were made to represent impact, both bibliometric, social and economic, map the effectiveness of resources gained from public bodies, and see which areas and policies most effectively converted investment into impact.

As a counterpoint to this, and not to be applied to all units, the anuário ought to find a way to represent private funding captured in for innovation, as previously mentioned, this is by no means a complete indicator for innovation or quality, but it would help us to understand better the distribution of private money in the institution. Public and private money are issues of a different order, and designed to measure different things. Public money is unquestionably a positive thing for any department or institution, while private is more ambiguous. It depends rather on what the private money is intended to do, and varies greatly depending on the nature of the agreement. However, industrial partnerships are important for areas like engineering and applied sciences, and so we should be able to see, across the institution, which departments are effectively working with the private sector, and how.

7.1.2 Section 5: Library Systems

One of the areas in which USP is a world leader is in the provision of open access online material. In the most recent Webometrics Repositories ranking, USP's registered in twelfth, the only other university in the top twenty is the Autonomous University of Barcelona. If, as stated in the earlier section on open access, the question is not if but when open access is universal, the university should be preparing and considering alternative methods of measurement in its open access activities. In the 2014 edition, the online measurement section has been substantially increased to reflect this. In prior editions, there was a representation of total downloads from the repository and a simple measure of the number of theses submitted broken down by unit.

In the 2014 edition the number of page unique views, hits and downloads are measured by website and the volume of work put up for open

access is measured. Through a combination of sections three and five it should be possible to give an idea of impact and performance of research as opposed to simple output, although it will be the contention of this section that these links are not made explicit enough, and there is a strong argument that the two sections should be more closely combined.

Furthermore, the simple reporting of hits and downloads of papers remains a blunt instrument to measure the effectiveness of social impact or health of a repository. While it is a representation of what leaves the repository, it is not a representation of how this knowledge is then used or transferred.

For all of the problems highlighted the beginning of the dissertation about measuring and thinking about impact, it is clearly this, and not raw production, productivity or cost/benefit that USP needs to measure. The state of São Paulo, and more specifically USP, diverges from the rest of Brazil sharply in research expenditure and in terms of output, sitting much closer to OECD averages than it does to the rest of Brazil. For this, the national CAPES evaluations, which focus on the need to increase output and implant an academic community in Brazil are not especially helpful for the university in measuring its performance. It is for this reason that international comparison and the adoption of more descriptive metrics is so important to the university, and also for this reason that the university should move its performance metrics away from national standards and towards engagement with the wider global discussions on how to measure performance. In this case analysis of Brazil's higher education system on a macro level brings misleading conclusions when considering the micro.

For USP, the results of staying with national standards of measurement have produced a situation in which productivity is valued above all other virtues, in deed if not in word. The number of articles published is indicative of quality in situations only where very little research is produced, which is undoubtedly true of certain parts of the country, but not within the state of São Paulo. The neo-liberal conception of the university as knowledge factory creates much ire within factions of the institution. In this critics are absolutely correct, that the metrics of productivity imported from the commercial world are totally unsuited to the process of evaluating quality or impact in higher education, because universities

fundamentally deal in post-capitalist goods, not in saleable units. Efficiency, cost/benefit and productivity will inevitably form part of any analysis, but they cannot form the *whole* analysis. Where critics of this process are wrong, however, is in the conclusion that we need less quantitative evaluation. For all the methodological and normative issues with current ranking systems, two salient issues arise from them; one, rankings are an inescapable fact of higher education today, and simply refusing to engage with them is unproductive, second, that in modern massified systems of education it is inconceivable that power of evaluation can be devolved to the academy again. We are confronted with a reality, and a normative duty. The university must respond to them, and the way in which this should be achieved is by utilising more measures in a more mature way than is currently seen in the university, or in rankings.

7.2 Conclusions

There are two important questions to be answered of the Anuário Estatístico; what is its actual purpose? Is the format and information contained within it fit for this purpose? One possible purpose is to present information publicly in order to contribute to discussion and transparency. If this is the case, then it is doubtful that in its current format, a 700 page pdf or volume composed of tables and simple graphs will ever attract much public debate or attention when compared to interactive university rankings. If it is intended to engage the public and contribute to discourse about higher education, then its format should be changed to something more user-friendly. Rankings have gained popularity because of their immediacy and intelligibility, this is not something with which we could accuse the Anuário Estatístico.

If it is to serve as a database of material for research on the university for researchers and policymakers to gain insight into the workings of the university, then it lacks the descriptive depth needed to represent impact or the university's place in world academia, and it lacks a format that allows for cross comparison with other sources. When it was founded, the amount of information or analysis of research possible was much lower, and the fact that the anuario had to be in physical form meant that the Anuário had to be recorded and

formatted in this way. Today the university has the ability to represent much more information in a much more intuitive way, but it must decide whether the Anuário is to be a research resource or a resource of public information, and structure it accordingly.

8. RECOMMENDATIONS AND FUTURE OPPORTUNITIES

Based on the results of the comparison between the rankings metrics and results and analysis of the Anuário Estatístico, this section will explore some alternative options the university could use to conceive of impact and quality that could go some way to resolving both the paradox of the need to measure and the impossibility of finding universal performance indicators and the need to compare and measure performance with the general undesirability of hierarchical rankings and league tables to measure institutions.

8.1 Fourth Wave Ranking Systems: Benchmarking

What is required to utilise these drivers towards transparency and comparability, then, is the idea put forward by the new 'fourth wave' of rankings (the first three being; pure reputational surveys, bibliometric rankings and composite rankings respectively). The fourth wave is characterised by being non-hierarchical, instead offering a broad range of metrics, which the user can then weight and value according to his or her priorities. This means that the 'ranking' is only ever produced at the end of the process and is not intended to demonstrate the all round 'best', but how institutions perform according to certain parameters (SALMI; 2013). This means that universities can be publicly compared without the power relations problem outlined above, and without being committed to any one view of impact or the role of a university. In 2012 the European Commission, upset by continual underperformance of European institutions in rankings as a result of the greater stratification of universities in the system, where there are very few recognisably 'comprehensive research-intensive' institutions of the type favoured by rankings, held a feasibility investigation on a project called U-Multirank, allowing users to individually assess institutions (Van Vugt & Ziegele; 2013), while the World Bank's benchmarking system focuses not on individual institutions but on 'systemic

health', encouraging the holistic development of HE systems and not pressuring governments into stratifying and using ranking position as justification for diverting resources into a few elite institutions at the expense of others.

An example of a benchmarking project that was presented to USP's reitoria in preparation for the USP80anos conference can be found in the appendix.

From the argument presented in the preceding chapters, we can see that normative prescriptions of a single vision of excellence will always be problematic for universities, and when analysed on a global level, not necessarily of enormous use for policymaking or better understanding of how universities work, what the likely outcome of change is, or how performance or excellence is defined. From this, what we ought to demand of our measurement tools are:

1. Tools that can select universities for comparison based on sensitivity to their geopolitical and national contexts and institutional characteristics, to better understand the effect of institutional change and policy.
2. Mechanisms that can represent impact in a variety of senses, depending on the context and area of the research, in a transparent way that can facilitate meaningful international comparison.
3. Mechanisms that can identify fruitful areas within disciplines, and suggest interface areas between them.

8.2 New Measurement Metrics

Establishing that the 'university of production' measurement used by the university is an unsatisfactory representation of performance of a university of USP's type, and the methodology used in international rankings measures performance a very constrained way that does not capture the diverse and fragmented nature of the modern research landscape, there is a clear dilemma for the future of measurement of research. We are faced with a choice of either abandoning the project of external evaluation of research performance, instead

reverting to the faculty level qualitative peer review characteristic of elite systems of higher education, or expanding and improving the way in which impact and performance is measured, depending less on simple productivity or citation impact measures and incorporating a variety of approaches.

This is an attempt to respond to the demand that the way in which we think about research performance should consist not only in expansion of what is measured, but also in a broadening of the means in which we undertake it. This section is exploratory, and is therefore intended to point out possibilities for future projects, rather than to actually carry them out.

With the enormous capabilities that technological progress have brought us, it is almost unfathomable that we would still use an algorithm written by Eugene Garfield in 1963 (DAS; 2015) designed for use in physical libraries and manual calculation. This algorithm works within the technical limits of what was possible in the 1960s, before digitalisation of information and before usage could be measured in any other way than physical citations. In 1963, a measure like an Eigenfactor would have required such a huge amount of information to manually process for each article; not just citations, but citations of cited works in a network, the measure would be infeasible. Furthermore, as all scholarly work was kept as a physical copy within a library, there was no way of measuring or assessing academic conversation outside of this. The vast expansion of computing power, and the way in which scholarly work is stored and distributed has presented an opportunity to remold the way in which we measure (MILOJEVIC; 2014).

Today, academic conversation takes place in such a variety of fora; from Twitter as preliminary conversations and new forms of peer review (DARLING, SHIFFMAN, COTE, DREW; forthcoming), to websites like Citeulike, Mendeley, ArXiv and Research Gate. It is the recording of these new sources, alongside non-academic ones, like newspapers, blogs and social media sites. It is these that Altmetrics attempt to measure.

This represents a significant 'opening up' of inputs in the measurement process (STIRLING, 2005); instead of the damagingly narrow number of inputs

that rankings utilise, there is acknowledgement that as academic behaviour changes over time, so too should the range and variety of materials used to assess impact. With this in mind, many research councils have been quick to begin to adopt alternative metrics (PIWOWAR; 2013), listing research products, as opposed to publications, be able to measure a variety of different media. Altmetrics are still imperfect measures of impact, and many academics have been quick to round on them (CHEUNG; 2013), for being easy to game, and being of questionable academic validity. This, I think, misses the intention of these new metrics, leaning on the old notion that they should be used to 'close down' and define the nature of quality or impact.

On the contrary, they in fact represent an 'broadening out' of the outputs in the research process. These alternative metrics do not seek to replace or reinforce the way in which metrics are used at present; to define and set the rules of the game, but instead to act as representative tools to foster more debate, more diversity in the way in which we measure, and more potential for shared learning geared towards institutionally specific goals, rather than the zero sum arms race towards the top of the rankings (RAFOLS et al; 2012, HAZELKORN; 2011).

Furthermore, Campbell's law (1973), which states that the more we rely on simplistic quantitative measures the greater the incentive for gaming those measures is diminished with the eradication of universal measures. Using a diversity of different measures will make gaming the system less attractive, as less is at stake with each individual measure.

It is not impact factor per se that causes issues, rather its crude adoption by policymakers, rankings and university management that is so problematic. This broadening of both input and output is incompatible with the construction of a hierarchical ranking, as it does not presume any specific hierarchy on types of publication, or types of impact. Rather, this is left to the judgment of the individual, or the individual institution.

We can use this type of analysis to other departments to highlight significant areas of interface, to facilitate the organisation of interdisciplinary working groups within the university. It could also be used to compare with other

departments in other university. This analysis is purely descriptive, and so cannot be used to measure performance to any universal standard, rather to diagnose whether the department is producing the kinds of work it intended to, and where significant areas of collaboration could be possible.

There are a number of standard file formats with which Scopus and WoS export to (.isi, for instance), allowing rapid access to bibliometric data and batch downloading of bibliometric data for analysis. At present, SciELO has limited compatibility; it does not record citation data on the WoS network, but does keep its information on the SciELO Citation index. It does follow standard ISI layout, meaning that publication data can be analysed next to international research. Citation information for products published outside of the WoS core collection is unavailable at present, it is unclear whether or not Thomson Reuters has plans to integrate the indexes or leave them separate, but for the time being the assessment cannot stretch beyond publication location.

The USP library system does not follow ISI format, nor does it record citation information. This makes the process of recording and mapping arduous, and often a manual exercise, like in the IRI map above. The university ought to record its bibliometric information in a recognised standard format, this would allow much easier comparison with sources like the WoS, and easier analysis.

8.2.1 Measurements of Impact

Scholarly communication should be conceived of as a series of overlaid networks of communication (MILOJEVIC:2014). We should be careful, therefore, when considering impact, not to give absolute privilege to one form of communication over all others.

Regular citation algorithms connect research in a linear way, from paper A (or journal A) to paper B (or journal B), meaning that all citations are given equal weighting, and the algorithm conceives of citations as isolated, not as a measurement of a part of a system or network. As mentioned earlier, article citation impact and journal citation impact are commonly used because they are

easy to use and understand, but are poor representations of true impact. Eigenfactors instead measure the relative centrality of a paper or journal relative to a map of science using the eigenvector formula outlined above.

At present, eigenfactors have been applied mainly to journal impact factor because they are first and foremost of interest to the academic community, which is more interested in maximising the reach of their published research through publishing in the most 'central' journal. Indeed, ISI have now integrated the Eigenfactor into the Web of Science citation report, but because it is a little difficult to interpret or understand when compared to citation rate or h-factor, its uptake has been somewhat slow. This method could easily be applied to institutions, seeking to represent their performance in a more detailed way.

8.4.5 Altmetrics

Whether or not Altmetrics ought to be called metrics at all is a subject of much debate at present (NEYLON; 2014). A metric for research is a criterion by which all research can be judged equally, compared and ranked. Altmetrics should be considered more as indicators; measures that can be indicative of performance, but should not be taken as externally valid metrics in and of themselves. For example, a ranking of the most popular articles on Facebook will not tell an assessor a great deal about the quality or necessarily even the outreach of a given article. A combination, however, of the number of readers on Mendeley, the number of times a paper has been cited in policy reports, news sites, academic blogs, reading lists and Twitter conversations could be indicative of a much wider application of the research than academic conversation. In this way we can measure what the social impact of a piece of research is.

The paper *Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000* (PALL et al; 2011) has a citation score of 179 on the Web of Science index. It is registered as highly cited in its field, but when compared to others from *Nature*, there is nothing immediately obvious about its impact. However, looking at the Altmetric score for this paper, we can

see that it appeared in 11 different news stories, 2 policy papers (one from Oxfam, one from the British Government), one wikipedia entry, it has 259 Mendeley readers, placing it in the 94th percentile for *Nature* articles. This means that it has contributed in a significant way to public education (in news sources, and in wikipedia), public discussion, as shown by its Twitter and blogging appearances, and in public policy formation. These dimensions should form an important part of our considerations of impact.

While Altmetrics cannot give us absolute measures of impact, they can trace access and interaction with a paper according to geographical location, which could be an important step in understanding the dynamic related to citation rates and geographical areas. This could allow better sense of the geographical global flow of citations than the current ISI or Scopus data, from which it is difficult to extract citation patterns from anything other than the location of the publisher, not the location of the researcher.

Altmetrics can also be implemented at pre-publication stages, for laboratory notebooks, conference proceedings, dissertations and any other digital research product. This means we can begin to track processes of collaborative online production and discussion, and the internationalisation not just of the published research, but also of its production and how well it engages with a variety of contexts and stakeholders both before and after its publication. This turns the process of opening access to information easier, because it is no longer necessary to publish in an indexed journal to have impact measured.

8.3 Concluding Remarks

Universities are, by their nature, multifaceted institutions that require a fundamentally different type of governance to other bodies. They are not information or workforce factories, and should not be measured as such. They are, furthermore, products of their history, environment and ideology. It is not clear that all universities can, or should, be 'ranked' according to some external, universal metric into a global list.

Rankings are not transparent, objective indicators of excellence that accurately measure change in institutions year on year. They are the reflection of power structures inherent in global research. In some cases, they are the product of powerful commercial interests, in others a reflection of specific state development goals. The use of a range of indicator types as a compound score raises serious issues about what is actually measured by rankings, as it is not a score of pure output or impact. The annual cycle is too short to properly reflect institutional change, and there are serious misgivings about the role of halo effects, statistical noise and gaming in the distribution of ranking position. Although there have been a number of attempts to address the lack of focus on teaching, ranking agencies have yet to come up with a stable or relevant base from which to compare teaching performance across national contexts. The neglect of social inclusion, extension and teaching in global rankings leads to heavy incentives away from a number of vital roles the university performs and into a relatively narrow band of elite research in exact sciences. Finally, they do not represent the diversity inherent in industrial missions adequately, nor do they represent a university's social contribution.

They do, however, shed some light into performance of institutions, and create a powerful public discourse to which universities are increasingly being held to account by the public. Because of this double bind, the university has no choice but to find some response to rankings, either competing for position and playing the game, or finding plausible alternatives.

Research and teaching can no longer be thought of purely in terms of academic impact as an end in itself, but as powerful drivers of societies and economies with the potential to mould future generations and reshape the nature of society. Because so much research is produced today in the context of its application in response to societal demand, research can have a huge variety of outcomes and impacts which cannot be assessed *ex ante* according to traditional metrics. Approaches to the big problems facing society are not the same as disciplinary scientific programs, which seek to either solve or pose cognitive questions to scientific paradigms.

Rankings have largely proliferated not because of their validity, or their ability to accurately represent the nature of a research institution. They have proliferated because they reduce extremely complex phenomena into easily digestible information and a format recognisable to non-practitioners. CAPES evaluations, similarly, hold such sway over Brazilian HE in part because they dictate resources allocation, but also because of their ability to give clear and objective scores over something that is very difficult to represent. Paradoxically, these calls are both impossible to ignore and undesirable to blindly satisfy for a modern institution.

To respond to this challenge, USP could find stable bases for comparison with institutions of a similar profile, and a robust but flexible range of indicators that can represent activity and impact within the university. If the university of the future must be flexible enough to respond to new possibilities, challenges and demands laid before it by society, it must ensure that its methods of measurement are flexible and agile enough to follow, accurately represent and reward it. There is an ongoing problem with USP, in that because of its heavily bureaucratic nature and federalist structure each unit is highly autonomous. This means that it is not well set up for change, or to benefit from top-down initiatives. If these competences exist within the university, it has yet to find a way to harness them, a channel to turn information gathered about the university into sound policy. This is something a professionalised Anuário Estatístico with a proper feedback mechanism into the decision-making process could assist.

The examples set out in the final part of this paper are just a small selection of the vast array of possibilities that the digitalisation of information presents us with, and we should take this opportunity to consider that, instead of closing down the way in which we measure research with metrics and rankings, we ought to open it out with representative indicators and benchmarking projects. This will make governance more capable of being reflexive with regards to its place in society and global higher education, and bring greater understanding and potential for shared learning with other institutions. These new data will be much more complex than rankings and more difficult to understand as a result, but the academic community must decide whether it is preferable to assess universities based on what is easy to measure, or what is genuinely valuable.

To satisfy the public demand for rankings, the university could present this information in a more accessible format than the *Anuário Estatístico* in order to open a public dialogue over what the university ought to strive towards. In this way, the often neglected third mission of universities can be allied more closely with the research mission, by better explaining and educating the public about what research does, and how it is performing. Public dialogue and a diversification of stakeholders has become a key aspect of massified higher education, with this comes a heightened need to communicate effectively with the public.

Rankings have sprung up largely in the absence of this effective communication, and therefore, if the university wishes to wrestle some of the initiative back from public ranking agencies then it must communicate in a way that is as engaging as rankings' easy conclusions and sweeping global generalisations that appear with each annual publication. Resolving Schwartzman's paradox (1993) of science becoming simultaneously more complex and more accessible and widespread within society will be a monumental challenge, but an essential one for the university in the future.

9. BIBLIOGRAPHY

ACADEMIC RANKINGS OF WORLD UNIVERSITIES. **Academic Ranking of World Universities**. Disponível em: <<http://www.shanghairanking.com/>>.

ALTBACH, R; SALMI, J (2011) **The Road to Academic Excellence : The Making of World-Class Research Universities**. World Bank (2011) p. 14

AXT, Margarete. **O pesquisador frente a avaliação na Pós-graduação: em pauta novos modos de subjetivação**. Psicologia e Sociedade (Impresso), Porto Alegre, v. 16, n.1, p. 69-85, 2004.

APPADURAI, A. **Modernity at Large**. Minnesota: University of Minnesota Press, 1996.

ARCHIBUGI, D. **Cosmopolitan Democracy and its Critics: A Review** European Journal of International Relations, 2004.

AROCENA, R.; SUTZ, J. Changing knowledge production and Latin American universities. **Research Policy**, v. 30, p. 1221–1234, 2001.

AROCENA, R.; SUTZ, J. **Latin American Universities: From an original revolution to an uncertain transition** Higher Education, 2005.

BALL, S. J. Privatising education, privatising education policy, privatising educational research: network governance and the “competition state.” **Journal of Education Policy**, v. 24, n. 1, p. 83–99, jan. 2009.

BARTUNEK, J. M. **What has happened to mode 2?** **British Journal of Management**, 2011.

BATY, P. “THE World University Rankings.” **The Times Higher Education Supplement**, 2010.

BENNET, D et al. (2007) **Open Letter to College Presidents** available from http://www.educationconservancy.org/presidents_letter.html

BERNHEIM, C. T. (2008) **La educación superior en América Latina y el Caribe : diez años después de la Conferencia Mundial de 1998**. IESALC, Colombia (2008).

BJÖRK, B.-C. et al. Anatomy of green open access. **Journal of the Association for Information Science and Technology**, v. 65, n. 2, p. 237–250, 6 fev. 2014.

BOOKSTEIN, F. L. et al. Too much noise in the Times Higher Education rankings. **Scientometrics**, v. 85, p. 295–299, 2010.

BORNMANN, L. **Measuring the broader impact of research: The potential of altmetrics**. 27 jun. 2014.

BOWMAN, N. A.; BASTEDO, M. N. Anchoring effects in world university rankings: Exploring biases in reputation scores. **Higher Education**, v. 61, p. 431–444, 2011.

BREMBS, B.; BUTTON, K.; MUNAFÒ, M. Deep impact: unintended consequences of journal rank. **Frontiers in human neuroscience**, v. 7, p. 291, 23 jan. 2013.

BONACCORSI, A (2008) **Search Regimes and Industrial Dynamics of Science** in *Minerva* vol. 46 pp285-315

CAÑIBANO, C; OTAMENDI, F; SOLÍS, F (2011) **International Temporary Mobility of Researchers, a Cross-disciplinary Study** in *Scientometrics* vol.89 no. 2

CAMPBELL, D. T. (1976) **Assessing the Impact of Planned Social Change**, Paper No. 8 Occasional Paper Series

CHAIMOVICH, H (2000) **Brasil, ciência, tecnologia: alguns dilemas e desafios** Estudos avançados. vol.14 no.40 São Paulo Sept./Dec. 2000

CHENG, Y.; LIU, N. C. A first approach to the classification of the top 500 world universities by their disciplinary characteristics using scientometrics. **Scientometrics**, v. 68, n. 1, p. 135–150cheng, 2006.

CHEUNG, M.K (2013) **Altmetrics; too soon for use in assessment** in *Nature* vol. 494 p176

COHEN, S.; FOUCAULT, M. **Discipline and Punish: The Birth of the Prison.***Contemporary Sociology*, 1978.

COLYVAS, J. **Performance Metrics as Formal Structures and through the Lens of Social Mechanisms : When Do They Work and How Do They Influence ?** v. 118, n. 2, p. 167–197, 2014.

CROSSLEY *, M.; TIKLY, L. **Postcolonial perspectives and comparative and international research in education: a critical introduction***Comparative Education*, 2004.

DARLING, E; SHIFFMAN, D; COTE, I; DREW, J (forthcoming) **The role of twitter in the life cycle of a scientific publication**, forthcoming, available via <http://arxiv.org/pdf/1305.0435v1.pdf>)

DAS, A. K (2015) **Article and Author Level Measurements** in Mishra and Sen eds. *Open Access for Researchers, Module 4: Research Evaluation Metrics*. Paris: UNESCO, pp. 61-91.

DOCAMPO, D. Reproducibility of the Shanghai academic ranking of world universities results. **Scientometrics**, 2012.

ERNE, R. **on the use and abuse of bibliometric performance indicators: a**

- critique of Hix's "global ranking of political science departments."** *European Political Science*, v. 6, n. 3, p. 306–314, 1 set. 2007.
- ETZKOWITZ, H; LEYERSDORFF, L: **Dynamics of Innovation: From National Systems and Mode 2 to a Triple Helix of university-industry-government relations**, *Policy Research* 29 (2000) 109-123, Elsevier
- FORGACS, D. **THE GRAMSCI READER Selected Writings 1916-1935** edited by. Forgacs, D, NYU Press (2000).
- FRENKEN, K; HOEKMAN (2014) Spatial Scientometrics and Scholarly Impact in Ding eds. **Measuring Scholarly Impact: Methods and Practice**, Springer, 2014
- GAUCHAT, G. Politicization of Science in the Public Sphere: A Study of Public Trust in the United States, 1974 to 2010. **American Sociological Review**, v. 77, n. 2, p. 167–187, 29 mar. 2012.
- GIACALONE, R. A. **Academic Rankings in Research Institutions: A Case of Skewed Mind-Sets and Professional Amnesia** *Academy of Management Learning & Education*, 2009.
- GINGRAS, Y. **How to boost your university up the rankings - University World News**. Disponível em: <<http://www.universityworldnews.com/article.php?story=20140715142345754>>. Acesso em: 21 ago. 2014.
- GLADWELL, M (2011) **The Order of Things** in *New Yorker*, February 14 2011 <http://www.newyorker.com/magazine/2011/02/14/the-order-of-things>
- GLÄNZEL, W. Bibliometrics as a research field. **Techniques**, v. 20, p. 2005, 2003.
- GOLDEMBERG, J (1996) **A USP as Décadas** in *Revista ADUSP*, Dezembro 1996 <http://www.adusp.org.br/files/revistas/08/r08a04.pdf>
- HABERMAS, J. The Idea of the University - Learning Processes. **New German Critique**, n. 41, p. 3–22, 1987.
- HALLIDAY, F. The chimera of the " International University ." **International Affairs**, v. 75, p. 99–120, 1999.
- HARNAD, S. et al. The Access/Impact Problem and the Green and Gold Roads to Open Access: An Update. **Serials Review**, v. 34, p. 36–40, 2008.
- HAZELKORN, E. "Problem with University Rankings." **Science and Development Network: Science and Innovation Policy Aid for Higher Education**, 2009.
- HAZELKORN, E. (2011) How Rankings are Reshaping Higher Education. In: **Rankings and the Reshaping of Higher Education: The Battle for World-Class Excellence**. . Palgrave, London, 2011. p. 1 – 8.

HICKS, D. The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. **Scientometrics**, v. 44, n. 2, p. 193–215, fev. 1999.

HOFFMANN, C. P.; LUTZ, C.; MECKEL, M. **Impact Factor 2.0: Applying Social Network Analysis to Scientific Impact Assessment** 2014 47th Hawaii International Conference on System Sciences. **Anais...IEEE**, jan. 2014

HOHENDAHL, P. U. Humboldt Revisited: Liberal Education, University Reform, and the Opposition to the Neoliberal University. **New German Critique**, v. 38, n. 2 113, p. 159–196, 22 jul. 2011.

HUANG, M. H. Opening the black box of QS world university rankings. **Research Evaluation**, v. 21, p. 71–78, 2012.

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE (2007) **Charter and Statutes 2007** available via [https://www.imperial.ac.uk/media/imperial-college/administration-and-support-services/secretariat/public/college-governance/charters-statutes-ordinances-regulations/charters-and-statutes/CHARTER---July-2007--\(Rev--Nov-2014\).pdf](https://www.imperial.ac.uk/media/imperial-college/administration-and-support-services/secretariat/public/college-governance/charters-statutes-ordinances-regulations/charters-and-statutes/CHARTER---July-2007--(Rev--Nov-2014).pdf)

INTERNATIONAL LABOUR ASSOCIATION; **Resolution Concerning Updating the International Standard Classification of Occupations** (2007) available from <http://www.ilo.org/public/english/bureau/stat/isco/docs/resol08.pdf>

KIBRE, P. Scholarly Privileges : Their Roman Origins and Medieval Expression. **The American Historical Review**, v. 59, n. 3, p. 543–567, 1954.

KUHN, T. **The Structure of Scientific Revolutions (1962)**, University of Chicago Press . p. 9–66

LAUDAN, L (1977) **Progress and its Problems; Towards a Theory of Scientific Growth**, University of California Press (1978) pp70-121

LATOUR, B. **Science in Action**. 1987, University of Minnesota Press

LAWSON, C (2013) Making interdisciplinary research work – A review of the literature **Science, Technology and Society Priority Group** University of Nottingham Research Group <https://nottingham.ac.uk/sciencetechnologyandsociety/documents/stspg-interdis-report.pdf>

LEITE, D. **Brazilian higher education from a post - colonial perspective**. n. September 2012, p. 37–41, 2010.

LETA, J.; PEREIRA, J. C. R.; CHAIMOVICH, H. The Life Sciences - the relative contribution of the University of São Paulo to the highest impact factor journals and to those with the largest number of articles, 1980 to 1999. **Scientometrics**, v. 63, n. 3, p. 599–616, jun. 2005.

LEVITT, J. M.; THELWALL, M. Is multidisciplinary research more highly cited?

- A macrolevel study. **Journal of the American Society for Information Science and Technology**, v. 59, n. 12, p. 1973–1984, out. 2008.
- LEYDESDORFF, L. Caveats for the Use of Citation Indicators in Research and Journal Evaluations. p. 278–287, 7 nov. 2009.
- LINKLATER, A. **Cosmopolitan citizenship; Citizenship Studies**, 1998.
- LO, W. Y. W. Soft power, university rankings and knowledge production: distinctions between hegemony and self-determination in higher education. **Comparative Education**, v. 47, n. 2, p. 209–222, maio 2011.
- LORENZI, M. Power: A Radical View by Steven Lukes. **Crossroads**, v. 6, p. 87–95, 2005.
- LUHMANN, N (1995) **Social Systems**, trans. Bednarz, Stamford University Press, Stamford, 2003
- MANE, K; BÖRNER, K (2004) Mapping Topics and Topic Bursts in Proceedings of the National Academy of Sciences of the United States of America access via http://www.pnas.org/content/101/suppl_1/5287.abstract
- MARGINSON, S. **Open Source Knowledge and University Rankings Thesis Eleven**, 2009.
- MARGINSON, S. National and international rankings of higher education. In: **International Encyclopedia of Education**.. p. 546–553.
- MARTIN, B. R.; IRVINE, J. Assessing basic research Some partial indicators. 1980.
- MARTIN, S; BROWN, W. M. Brown; BOYACK, K. "DrL: Distributed Recursive (Graph) Layout," in preparation for Journal of Graph Algorithms and Applications.
- MARTINS, L. L. **A Model of the Effects of Reputational Rankings on Organizational Change Organization Science**, 2005.
- MELLO, H; ALMEIDA, J.M.C, ETZKOVITZ, H; **Incubadora Tecnológica de Caxias do Sul: Inovação Tecnológica sob a Perspectiva da Hélice Tríplice**, working paper, 2005, available via <http://www.ucs.br/etc/conferencias/index.php/mostraucsp PGA/mostrappga2013/paper/viewFile/3611/1111>)
- MILOJEVIC, S (2014) Network Analysis and Indicators in Ding eds. **Measuring Scholarly Impact: Methods and Practice**, Springer, 2014
- MOREL, C; SERRUYA, S; PENNA, G, GUIMARÃES, R; (2009) **Co-authorship Network Analysis: A Powerful Tool for Strategic Planning of Research, Development and Capacity Building of Neglected Diseases** in PloS Neglected Tropical Diseases vol. 3 no. 8 available via <http://www.plosntds.org/article/fetchObject.action?uri=info:doi/10.1371/journal.pntd.0000501&representation=PDF>

MOURA, B.; AMORIM, V. **Para professores, queda da USP em ranking não significa menor produção acadêmica**. Disponível em: <<http://www.jornaldocampus.usp.br/index.php/2013/11/para-professores-the-nao-e-criterio/>>. Acesso em: 8 set. 2014.

NELSON, W.; SAUDER, M.; ESPELAND, W. N. Rankings and Reactivity : How Public Measures Recreate Social Worlds 1. v. 113, n. 1, p. 1–40, 2014.

NOWOTNY, H.; SCOTT, P.; GIBBONS, M. **Introduction: “Mode 2” revisited: The new production of knowledge** *Minerva*, 2003a.

NOWOTNY, H.; SCOTT, P.; GIBBONS, M. Introduction: `Mode 2' Revisited: The New Production of Knowledge. *Minerva*, v. 41, n. 3, p. 179–194, 1 set. 2003b.

NUNES, L; Interview granted to the author: April 2015

NYE, JR., J. S. Soft Power. *Foreign Policy*, p. 153–171, 1990.

O'LEARY, J.; INCE, M.; AL HILI, D.(2010) **QS World University Rankings**. Quacquarelli Symonds. p. 24 (<http://www.iu.qs.com/product/2010-qs-word-university-rankings-supplement/>)

OLSBO, P. Does Openness and Open Access Policy Relate to the Success of Universities? 2013.

OLSEN, J (2005) **Institutional Dynamics of a (European) Research Institution**, working paper Arena Centre for European Studies, University of Oslo available from https://www.sv.uio.no/arena/english/research/publications/arena-publications/workingpapers/working-papers2005/wp05_15.pdf

OPHTHOF, T.; LEYDESDORFF, L. **Caveats for the journal and field normalizations in the CWTS (“Leiden”) evaluations of research performance**. 14 fev. 2010.

ORDORIKA, I. El ranking Times en el mercado del prestigio universitario. *Perfiles educativos*, v. XXXII, p. 8–29, 2010.

ORDORIKA, I.; LLOYD, M. A Decade of International University Rankings: a Critical Perspective from Latin America. In: MAROPE, P. T. .; WELLS, P. .; HAZELKORN, E. (Eds.). **Rankings and Accountability in Higher Education Uses and Misuses**. .

ORDORIKA, I; PUSSER, B; **Bringing Political Theory to University Governance: A comparative analysis of Governing Boards at the Universidad Autonoma de Mexico and the University of California** in Smart eds. Higher Education Handbook of Theory and Research (2001) Agathon pp147-195

PALL, P; AINA, T; STONE, D.A; STOTT, P.A; NOZAWA, T; HILBERTS, AGJ;

LOHMAN, D; ALLEN, M.R; (2011) **Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000** in *Nature* vol.470

PHILLIPSON, R. (1992)**Linguistic imperialism**. OUP, Oxford. .

PIWOWAR, H (2013) **Value All Research Products** in *Nature* vol. 493 p159

POPPER, K (1959) **The Logic of Scientific Discovery**, Basic Books, New York, NY, 1959.

PUSSER, B.; MARGINSON, S. University Rankings in Critical Perspective. **The Journal of Higher Education**, v. 84, n. 4, p. 544–568, 2013.

RAFOLS, I. et al. (2012) How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management. **Research Policy**, v. 41, n. 7, p. 1262–1282, set. 2012.

RAFOLS, I; CIARLI; VAN ZWANENBERG, P and STIRLING, A (2012) **Towards indicators for ‘opening up’ science and technology**; policy proceedings of 17th International Conference on Science and Technology Indicators

RAFOLS, I (2014) Knowledge integration and diffusion: Measures and mapping of diversity and coherence in Ding eds. **Measuring Scholarly Impact: Methods and Practice**, Springer, 2014

RANGA, M and ETZKOWITZ, H (2013) **Triple Helix Systems: An Analytical Framework for Innovation Policy and Practice in the Knowledge Society**, in *Industry and Higher Education* 27 (4): 237-262

RANSON, S. Public accountability in the age of neo- liberal governance. **Journal of Education Policy**, v. 18, n. 5, p. 459–480, out. 2003.

RAUHVARGERS, A. (2011) **Global University Rankings and Their Impact**. European Universities Association Report on Rankings 2011

RIGHETTI, S (2013) USP Despenca no Principal Ranking Universitario da Atualidade in Folha de São Paulo, 02/10/2013
<http://www1.folha.uol.com.br/educacao/2013/10/1350715-usp-despenca-no-principal-ranking-universitario-da-atualidade.shtml>

RHOADES, G.; SLAUGHTER, S. Academic capitalism in the new economy: Challenges and choices. **American Federation of Teachers**, p. 37–59, 2010.

SAFÓN, V. What do global university rankings really measure? The search for the X factor and the X entity. **Scientometrics**, v. 97, p. 223–244, 2013.

SALMI, J. **The Challenge of Establishing World-Class Universities - ISBN:**

9780821378656. Washington D.C: The World Bank, 2009.

SALMI, J.; ALTBACH, P. G. **The Road to Academic Excellence The Making of World-Class Research Universities**. Washington D.C: The World Bank, 2011.

SCHUMPETER, J: **Capitalism, Socialism and Democracy** (1943) Routledge, New York (2003) p81

SCHWARTZMAN, S. OS PARADOXOS DA CIÊNCIA E DA TECNOLOGIA *. **Ciencia Hoje**, n. 95, p. 28–35, 1993.

SCHWARTZMAN, S. **Equity, quality and relevance in higher education in Brazil Anais da Academia Brasileira de Ciências**, 2004.

SCHWARTZMAN, S. Brazil's leading university: between intelligentsia, world standards and social inclusion. p. 1–36, 2005.

SEGERSTRALE, U (2000) **Beyond the Science Wars: the Missing Discourse about Science and Society**, SUNY Press

SEGLIN, P. O. Why the impact factor of journals should not be used for evaluating research. **BMJ**, v. 314, n. 7079, p. 497–497, 15 fev. 1997.

SLAUGHTER, S.; LESLIE, L. L. **Academic Capitalism: Politics, Policies, and the Entrepreneurial University**.. p. 296

SMITH, N (1987) **Academic War over the field of Geography: The Elimination of Geography at Harvard 1947-1951** in *Annals of the Association of American Geographers* vol.77 no.2 1987

STIGLITZ, J. **Knowledge as a Global Public Good**. In: KAUL, I.; GRUNBERG, I.; STERN, M. (Eds.). **Global Public Goods**. New York: UNDP, 1999. p. 308.

STIRLING, A (2007) **A general framework for analysing diversity in science, technology and society** in *Journal of the Royal Society Interface* vol. 4 (2007) available at <http://rsif.royalsocietypublishing.org/content/royinterface/4/15/707.full.pdf>

STIRLING, A. 2005. **Opening Up or Closing Down: analysis, participation and power in the social appraisal of technology**, in M. Leach, I. Scoones, B. Wynne, (eds) 'Science and citizens: globalization and the challenge of engagement', Zed, London, 2005, 218-231.

TALBURT, S. **Academic Capitalism and the New Economy: Markets, State, and Higher Education (review)**The *Review of Higher Education*, 2005.

THELEN, K (2012) **Varieties of Capitalism: Trajectories of Liberalization and the New Politics of Social Solidarity** in the *Annual of American Political Science* vol. 15 pp137-159

THELWALL, M.; VAUGHAN, L.; BJORNEBORN, L. Webometrics. **ANNUAL REVIEW OF INFORMATION SCIENCE AND TECHNOLOGY**, v. 39, p. 81–135, 2005.

TIKLY *, L. **Education and the new imperialism**, *Comparative Education*, 2004.

TOGIA, A.; TSIGILIS, N. Impact factor and education journals: a critical examination and analysis. **International Journal of Educational Research**, v. 45, n. 6, p. 362–379, jan. 2006.

TROW, M. **Problems in the transition from elite to mass higher education**. 1973.

UNDP; **Human Development Index Report 2014** available at <http://hdr.undp.org/sites/default/files/hdr14-report-en-1.pdf>

UNESCO. **Higher Education in the Twenty-first Century Vision and Action**. 2001.

UNESCO **International Standard Classification of Education 2011**, available via <http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf>

UNESCO. **UNESCO SCIENCE REPORT 2010. Executive Summary The Current Status of Science around the World..** v. 52p. 541

UNESCO-CEPES **Berlin Principles on Ranking Higher Education Institutions** (2006)
http://www.che.de/downloads/Berlin_Principles_IREG_534.pdf

UNIVERSITY OF CALIFORNIA SENATE, **Open Access Policy for the Academic Senate of the University of California Adopted 7/24/2013** available via <http://osc.universityofcalifornia.edu/open-access-policy/policy-text/>

UNIVERSITY OF CAMBRIDGE (2009) **Appointment of Vice Chancellor at the University of Cambridge 2009**, Perrett Laver, access via <http://www.statslab.cam.ac.uk/~frank/12493b.pdf>

UNIVERSITY OF OXFORD (2002) Statutes available from <https://www.admin.ox.ac.uk/statutes/781-121.shtml>

UNIVERSIDADE DE SÃO PAULO. (2013) **Anuário Estatístico USP 2013**.

VAN RAAN, A. F. J.; VAN LEEUWEN, T. N.; VISSER, M. S. Severe language effect in university rankings: Particularly Germany and France are wronged in citation-based rankings. **Scientometrics**, v. 88, p. 495–498, 2011a.

VAN RAAN, T.; VAN LEEUWEN, T.; VISSER, M. **Non-English papers decrease rankings**. *Nature*, 2011b.

WANG, Q.; CHENG, Y.; LIU, N. C. **Building World-Class Universities**. Boston: Sense Publishers, 2012.

WELLEN, R. Taking on Commercial Scholarly Journals: Reflections on the “Open Access” Movement. **Journal of Academic Ethics**, v. 2, n. 1, p. 101–118, 2004.

WELLEN, R. Open Access, Megajournals, and MOOCs: On the Political Economy of Academic Unbundling. **SAGE Open**, v. 3, n. 4, p. 2158244013507271–, 23 out. 2013.

WHITT, J. A.; LUKES, S. **Power: A Radical View**. **Contemporary Sociology**, 1980.

YUE, W.P. and ZHU, J.W. (2009). A quantitative perspective on the leading role of top universities in China. In **3rd International Conference on World-Class Universities Proceedings**, page 109-119

ZUBIETA, A. F. (2009). **Recognition and weak ties: Is there a positive effect of postdoctoral position on academic performance and career development?** *Research Evaluation*, 18(2), 105–115.

APPENDIX A: INSTITUTIONAL BENCHMARKING PROJECT

MARCOVITCH, J; AXEL-BERG, J (2014)

This benchmarking exercise was produced as an example of how with better and more sensitive recording of information and properly selected comparison for metrics valid on an institutional level to be compared, whilst allowing individual comparison by area or department allowing greater flexibility and sensitivity towards heterogeneity between areas. Benchmarking seeks not to produce objective lists of best and worst, or to prescribe best methodology. Instead, its aim on an institutional level is to provide stable bases for comparison between universities, avoiding the aggregation problems inherent in rankings.

The report was produced by prof. Jacques Marcovitch and Justin Axel-Berg in December 2014, and an abridged version of this report was presented to the Reitoria of the Universidade de São Paulo during consultation for the USP 80 Anos round of debates and consultation on reform demonstrating the value of more focused, considered comparison of institutions over the adoption of a universal playing field and universal metrics. The universities selected are all regional leaders who play a significant role in the higher education and research landscape of their societies.

BASE SELECTION CRITERIA

All institutions on our benchmarking list share the following criteria:

1. Size. It is difficult to compare institutions of vastly different size, as there are both competitive advantages to being small and intensive, and to being large scale multi-campus universities. As such, comparison

between them can be misleading. For the purposes of this exercise, all of the shortlisted institutions are multi-campus, with a minimum student body size of 20,000.

2. **Comprehensiveness.** Universities with dramatically reduced or specialized curricula have fewer of the balancing/competing interests problems outlined in the first section of this paper. As such their programs can be much more intensively focused, but their management does not confront the same types of dilemma in resource allocation. Therefore, all of the shortlisted universities are comprehensive, with human, exact and applied sciences.
3. **Public.** Private institutions have significantly different relations to states than public ones; they are typically less dependent on state resources for funding, and perhaps more significantly are not tied to the same types of social mission of expansion, extension or social utility of research as public institutions. Their governance and management can also vary wildly. For this reason, only public institutions are considered.

Variable inclusion criteria

In the interest of producing a list specifically of use to USP, universities were then selected on the basis of a variety of criteria to show a diversity of information.

1. **National development level.** The state of a population at entry is a highly relevant factor when considering university comparison. The Human Development Index was chosen because it includes GNP per capita and years of schooling. As such it should provide a rough proxy for a population's readiness for higher education, and many of the institutions may face similar, although not identical issues. For this a list was produced using a number of institutions in the High Development category for the HDI, as well as from the categories above it. The project

intended to combine this with similar Gini coefficient scores, but the group did not yield any promising universities with which to compare as a result of Brazil's notably poor performance in it.

2. Linguistic and geopolitical diversity. As established by a number of empirical studies, not using the English language severely reduces the size of the audience, and therefore academic impact, of research. For this, a variety of linguistic backgrounds was selected to see if there would be a variation in strategy to overcome this hurdle. For geopolitical context, the list is chosen not simply to represent universities in the centre, or only on the periphery, but a mixture of the two to aid understanding about the different challenges faced by each.
3. Openness and access to information. As USP takes its commitment to the extension of research very seriously, some effort has been made to select universities with similarly active profiles in OA. As the U-Multirank project does not as of yet measure this, the Webometrics Ranking of Universities was used.
4. Age of the institution. Older institutions not only benefit from enhanced halo effects of prestige, but also typically have much more embedded systems of governance and organization, and are therefore typically less prone to institutional change, one of the main purposes of the conference for which this benchmarking project was prepared. Where newer institutions were strong candidates, they were preferred to older counterparts.

UNIVERSITIES SELECTED FOR SHORTLIST

Universidad Autónoma de Mexico (UNAM)

Peking University

University of Copenhagen

Université de Paris (PRES X)

University of California

Nanyang Technological University

University of Tokyo

Universidad Autónoma de Mexico (UNAM) was selected because it fits into the same HDI bracket as USP, and in terms of profile, is USP's nearest comparative university. It is the most prominent university in Mexico, large in scale, multi-campus and substantially similar in both mission and structure to USP. Peking University was chosen because like UNAM and the University of Tokyo it confronts challenges with the English language and sits far from the research 'centre'. Peking however, is erroneously excluded from the U-Multirank's description of being comprehensive because it does not have a school of biology, although it does have one of life sciences. This would appear to be exclusion by technicality not by design.

Nanyang Technological University was chosen as a regional leader over the University of Singapore because it is newer and engaged in an intense program of reform and technology transfer, and therefore ought to produce more interesting results for comparison.

The sole Anglo-American entry to the shortlist is the University of California, because its structure as an extensive multi-campus (8 campuses) spread over a large geographical region with a high level of devolved autonomy. As such it shares many characteristics with USP. In 2013, it also embarked on the most ambitious green open access project yet undertaken to provide all future research from the university free to the public (UNIVERSITY OF CALIFORNIA SENATE; 2013). Its repository ranks comparably with USP's on the Webometrics Repositories ranking, and would be one of the top ranked in the Webometrics ranking itself if its production were aggregated rather than represented individually by campus.

The University of Copenhagen was selected as a classic example of an elite European university, a model that USP is based, while Paris was chosen

because it is a new entity made up of two old and extremely venerable institutions; Marie et Pierre Curie and Paris Sorbonne. The merger has led to a series of challenges and new ideas in university governance, and as such should provide a rich source of information.

Looking at the ranking position of the universities shortlisted, it would be difficult to see why these universities would be selected rather than any others. Indeed, when we look at the relative performance of the eight across the three rankings analysed in this paper, plus the HEEACT Taiwanese rankings, the variation in position is striking. Only the University of Tokyo displays consistent performance in every ranking. In this case, rankings are not especially helpful for actually determining performance to select suitable comparisons.

	Shanghai Jiao Tong ARWU	Times Higher Education	Quacquarelli Symonds	HEEACT
USP	101-150	226-250	132	62
NTU	151-200	76	39	123
UC Berkeley/UCLA	4/12	8/12	27/37	8/6
Paris (Sorbonne/Marie et Pierre Curie)	NA/35	DNR/103	227/115	NA/43
Peking	101-150	45	57	66
Tokyo	21	23	31	19
Copenhagen	39	160	45	30
UNAM	201-300	DNR	175	235

Basic information was gathered on each university where available from online sources and arranged into the table below. The information is not complete in all cases, as there are huge asymmetries in both the content and the way in which data are recorded across universities. This is probably due to differing higher education systems, differing priorities and in the case of financial data, different funding systems. The way in which Paris lists campuses, for instance, includes a large number of small institutes and

research nuclei scattered across the city, which is why there are 29 listed campuses; most universities would not categorise them as such.

Universities Comparative Table (information accurate to 2013)

UNIVERSITIES	Universidade de São Paulo	Nanyang Technological University	University of Califórnia	Universidade Autônoma do México	University of Copenhagen	Peking University	University of Tokyo	Université Paris-Sorbonne
Undergraduates	58.303	23.500	183.061	302.689 (Lic. e bacharelado)	23.473	15.128	14.120	40.000
Post Graduates	20.498	10.000	40.298	26.878	17.393	15.119	13.878	18.000
Total Students	96.701	33.500	233.198	330.382	40.866	30.247	28.133	58.000
Number of campuses	7	3	10	13 (com centros de extensão no exterior)	4	2	2	29
Resources from the state	3.986.741,82 (≈ US\$ 2 billion)*	US\$787m (55% budget: US\$1.47bn)	US\$2.38bi (budget: US\$24.1bn)	US\$2.47bi	US\$1.39bi	Unavailable	US\$800mi	US\$856m
Places offered annually	10.602		40.000	53.700			4.475	
Number of graduating students from UG	7.665	5916	46.953	45.544	7.663	6.959	3664	
Titles granted (masters)	6.016	2.928 (PG total)		7.347 (PG total)	3.604	3.449	2.958	
Titles granted (doctorate)	3.577		8.422 (PG combined)		693	851	1.153	
Publications (national)	18.390			3.994	-			-
Publications (international)	8.314	11.076		3.018	18.782	17.330	30.047	-

Academic staff	4.755	4.005 (2.333 are researchers - definition is not made clear)	16.300	5.478 permanents from 37.610 total	5.023	4.216	3.875	7.685
Non academic staff	14.589	2.834	133.000	28.400	4.249	4.422	3.752	7.830
Number of applicants through F U V E S T or equivalent	146.845	N/A	130.000		-		12.237	-
Mean spend per student	41.227 (≈ US\$ 20 mil)	US\$23.500	US\$6.100 (do US\$16.53 0 total)	US\$7.496	US\$34mil		US\$28.455	US\$14.758
UG Graduation rate (total UG students/graduating students)	7.61	2.66	3.9	6.65	3.06	2.17	1.83	
PG Graduation rate (expressed as % of enrolled undergraduates receiving degrees)	13.14	25.17	25.64	15.04	32.64	22.79	25.94	
Total number of alumni	≈ 300 mil	179.823	1.7M		-			-

Nanyang Technological University 2013: (http://www.ntu.edu.sg/bak_oas/Documents/NTU%20at%20a%20Glance%202013.pdf)

University of California: (<http://accountability.universityofcalifornia.edu/documents/accountabilityreport13.pdf>)

Universidad Autónoma de México: (<http://www.planeacion.unam.mx/Agenda/2013/pdf/Agenda2013.pdf>)

University of Copenhagen: (http://introduction.ku.dk/facts_and_figures/)

Peking University: (<http://www.oir.pku.edu.cn/En/html/AboutUs/AboutPKU/QuickFacts/>)

University of Tokyo (<http://www.u-tokyo.ac.jp/en/about/index.html>)

University of California- Berkeley: (<http://opa.berkeley.edu/campus-statistics>)

*Approximately 5% of ICMS – This does not include resources received from FAPESP, CAPES, CNPq or private initiative, for which there are no solid data. It refers only to what is received directly from the state. The cost per student only refers to the cost to the state.

**This PRES (alliance) has been operational since 2010; it is still in the process of integration. Therefore efforts to gather together all academic information has still not been carried out or published.

According to March and Olsen (2005), one of the things we can gain most from rankings and comparison is to understand the likely effects of policy and structural change to a university, so the broad structure, especially of the selection of executive, was carried out on our shortlist of 7:

USP

As defined in (Resolução no.6637, 2 October 2013)

Rectors are elected by the state governor for a fixed term of four years chosen from a list of three full time members of academic staff. These three are determined by an open election from the University Assembly composed of the University Council, central councils, groups from each unit and councils from museums. This assembly is composed of approximately 2000 representatives of the university community. The rector has autonomy in how he spends funding, but much less in how it is captured, he requires widespread support from a number of factions within the university to maintain power and is therefore short term goals or those benefitting the university community are often prioritized over the pursuit of long term or social and external goals. University autonomy is strong among individual academics and voting blocs, so representation is high. Management autonomy is significantly constrained both by dependence on state money, and therefore state assessment programmes, and to the satisfaction of internal interests as a result of the election system.

Peking University

As China has moved from a heavily coordinated centralised economy with a high demand for labour to a more liberal market economy with a high demand for technology over the course of the late 1990s, its universities have reformed significantly to answer to this changing demand. Prior to the 1998 China's universities were under direct control of the Communist Party of China (CPC); they were government-affiliated institutions with very limited autonomy whose principal aim was to supply qualified labour to state industry. The twin drivers of globalisation and a more diversified economy led to their universities becoming uncompetitive and too constrained to produce the technological innovation that the economy itself required.

Following the Education Law 1998 the Chinese government abandoned the direct control model in favour of a dual level system of control, where the CPC maintains overall control of the university through a representative council that convenes every five years to agree on development strategy and budget confirmation. Otherwise, the party holds a supervisory role over university governance and works in recruiting and training party members. It does not intervene in academic or day-to-day matters.

The managing group takes care of the university on a day-to-day basis, for maintenance and low level steering. It is composed of 8 people; the president and his vice presidents representing each faculty. There are no elections or representation mechanisms within Peking University's administration; all presidents and senior management are appointed by government agencies, and as such represent a significant constraint on both representation and academic freedom, especially in the social sciences and humanities, which are subject to party and political preferences, the prominence of the study of the writings of Mao Zedong being a particularly illustrative example. Because administrative power dominates over academic life, there is a real problem with developing truly world class research in Chinese universities.

In Peking then, inasmuch as it exists, institutional autonomy is granted not by any particular ideological or axiomatic belief, but because it is a more expedient method of achieving state goals. State goals and institutional goals are indistinguishable in China, because state and university are so intensely enmeshed.

UNAM

UNAM's history and governance structure are markedly similar to that of USP. Since 1944, the university has been presided over by a governing board, intended to de-politicise the institution and make final judgments on internal matters (PUSSER & ORDORIKI; 2001). The board is elected from the university council, approximately the same proportional size as USP's,

composed of the rector, institutional leaders and representatives from the teaching and student bodies. Board members sit for a fixed term of four years, renewable once to preside over all matters related to academic and administrative life. Differently to USP, it is the governing board which selects rectors of the university, rather than the university council (Art.6 Ley Orgánica de la UNAM; 1977 <http://www.dgelu.unam.mx/m2.htm>), which means that while student representation is still proportionally higher than in the other universities here, it is still lower than USP as it is at least mitigated by the separation between elections of the board who then appoint the rector.

UNIVERSITY OF CALIFORNIA

As an original land grant university, the university of California aggressively attempted to preserve autonomy from its benefactors in its early years. The university is governed through a board of Regents who oversee strategy and structure, and are appointed by the governor of California from illustrious members of the Californian society, not necessarily academics. Today, a regency is often granted as a political favour in return for support or donations during a gubernatorial election campaign rather than any specific expertise or knowledge of higher education. Regents are predominantly engaged in ensuring fiscal responsibility and in appointing the president to oversee more academic aspects of the university.

Participative democracy is low, with students and faculty typically under-represented in high level decision making processes. There is a student and faculty representative on the senate, but as the senate is a broadly deliberative not executive body, we can assess the University of California as having extremely high levels of institutional autonomy, as it is predominantly dependent upon its huge endowment fund (see chart), and not especially vulnerable to internal bargaining processes or state coercion.

UNIVERSITY OF COPENHAGEN

Legislative Decree no. 695 of 22 June 2011 on Universities (University Act)

<http://bestyrelse.ku.dk/vedtaegt/>

The University of Copenhagen has a governing board composed of 11 members, six of which are external appointments based on expertise in economics or academia, two of which are elected by the whole university staff, and two members elected by students. This changed in 2011 from an earlier, much more representative system of governance. The governing board is the supreme body within the university, and appoints the rector and oversees all administrative appointments.

The university is substantially autonomous from the state, but has a development contract developed in tandem with the ministry for science and technology that guarantees long term levels of investment for the fulfillment of basic institutional targets.

UNIVERSITY OF PARIS-SORBONNE

Universités Sorbonne have specifically separated their governance into an overall administration council, composed of the presidents of each constitutive part of the partnership, which have each been nominated by their respective schools to oversee funding allocation decisions in the short term. The academic senate composed of representative professors, researchers and students from within each school and from the broader population makes all decisions related to specific research programmes and curricula, and finally the strategic orientation committee, which is composed of experts from outside the university to oversee the long term goals of the university and assess how to meet them.

This model separates academic decisions from financial ones, depoliticising, to a certain extent, financial decisions while maintaining the connection between academic steering and academics. It is profoundly different in character from the model seen, for instance, in the University of California in its levels of representation and style of governance. This comes from a clear tradition and demand for academic autonomy in France, where state steering of

research is regarded with profound suspicion. It has, however, led to a series of crises in recent years (SALMI; 2009) where it appears that French higher education is not as competitive as the British or American institutions in rankings, and it has been suggested that it is specifically because of this inability to link financial issues to research steering, so it is more difficult to coordinate large projects.

UNIVERSITY OF TOKYO

The University of Tokyo Rules on Basic Organizations April 2004 <http://www.u-tokyo.ac.jp/content/400030172.pdf>

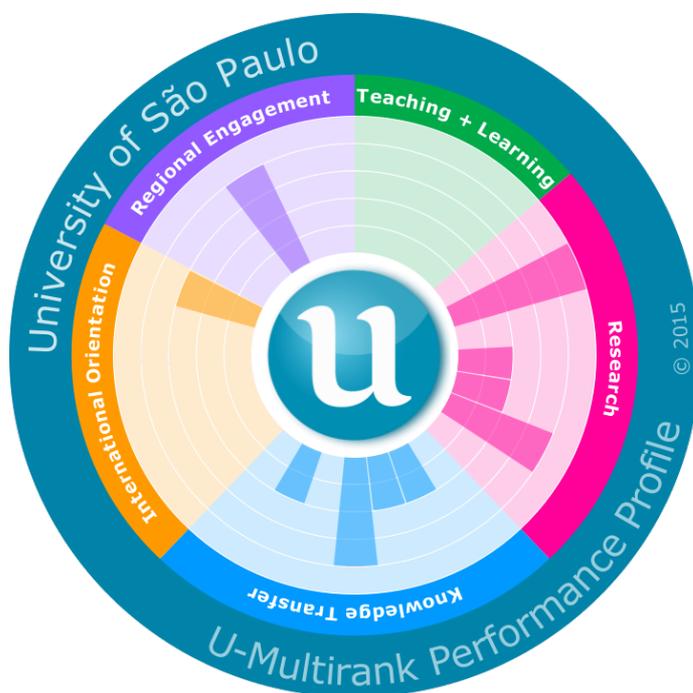
The national universities of Japan underwent a significant semi-privatisation in 2004, with all state institutions being turned into National University Corporations, with a view to substantially devolving power and autonomy to the institution itself. The president in Todai is selected by two councils, the administrative and academic boards. However, the organisation of the university is heavily influenced by private sector orthodoxy, and these two councils are largely advisory in nature. The executive is formed of the Office of the President and Council of the President, and he is in turn answerable to the board of trustees, not to the university itself, but the university corporation, which is defined as a separate legal entity.

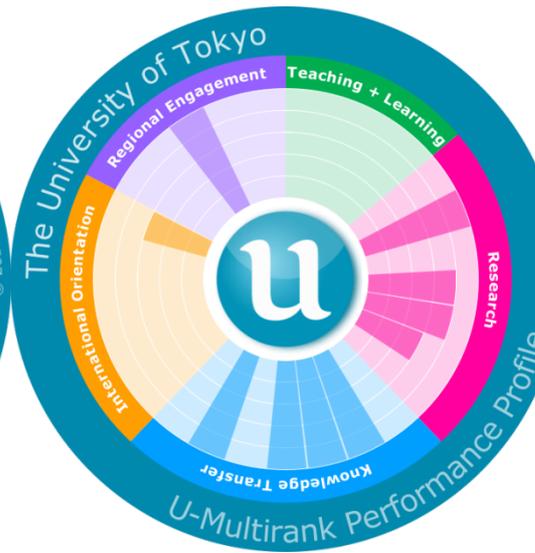
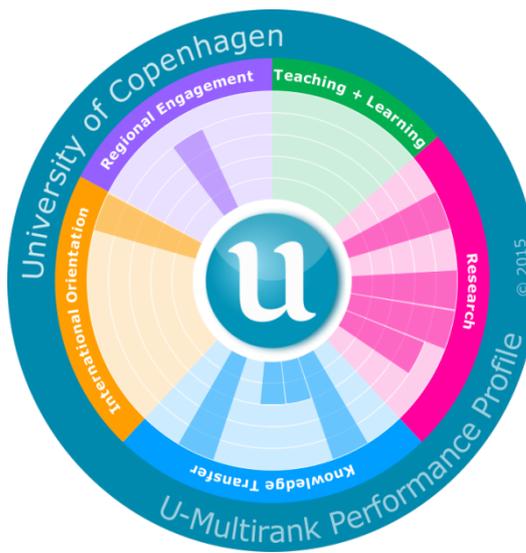
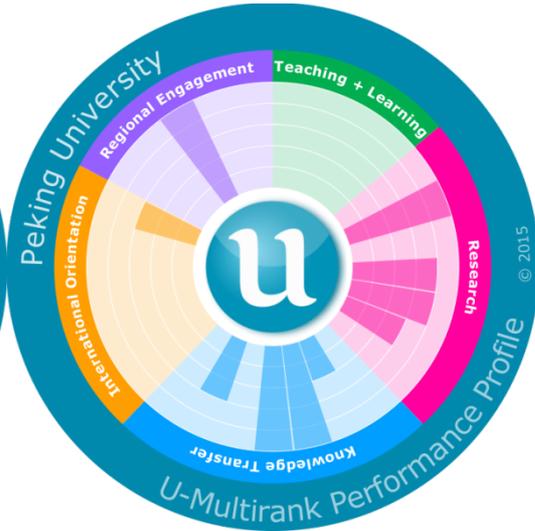
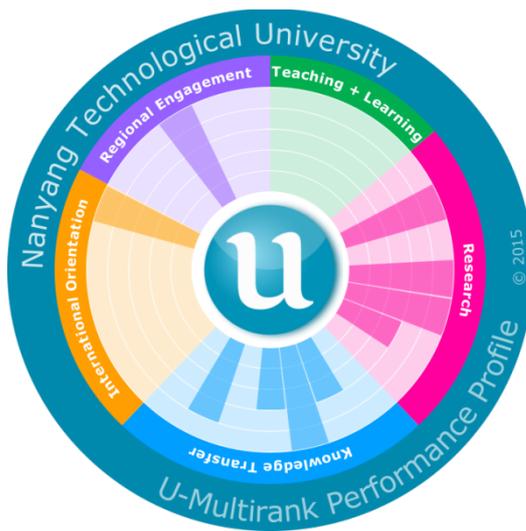
NANYANG TECHNOLOGICAL UNIVERSITY

As Nanyang is a relatively young institution, like many other younger Asian institutions it decided to base its management structure on the US model where the president is appointed through a board of governors, so as with the University of California, there is very little democratic participation in decision making, and institutional autonomy and steering is very high and state intervention is relatively low.

From these seven selected for their similarities to USP, although many share common characteristics, there is a huge variation in the types of management, organisational structure, levels of representation and accountability, each with its own unique relationship with its host state, from the very distant, like the University of California, to the intensively steered and controlled, like Peking University. If the aim of comparing universities is to better understand the likely outcome of change, and a better understanding of what results differing governance practices produce, then being able to benchmark according to these factors is vital.

With broad outlines of management structures established, the next stage of analysis involves looking at metrics for which it makes sense to compare on an institutional level. Below are the U-Multirank graphics for each institution, although it should be remembered that the project is still in its infancy, and there is still a lot of information missing from institutional profiles.







Peking University performs comparatively very strongly on patents awarded to researchers in absolute (45) and size normalised per thousand students (1.488) scores, and also on industry co-patents (40%), but not at all on co-publications with industry, or spinoffs, reflecting the particular triple helix configuration in China, wherein the emphasis is the state direction of innovation through its universities and into largely state run industrial partners. Tokyo, by comparison, has a much greater balance across the metrics, and a much stronger industrial profile than Peking with a focus on co-publications (7.425%)

The University of California ranks comparatively very low in terms of patents awarded to practicing academics, and very highly for co-publications. This is because it has long been more orthodox in the US not to develop technology itself in-house, but to develop the applied research (or Mode 2), which is then taken by the private sector and used in patents and for the innovation process itself. UNAM's industrial profile is extremely small, with one international patent granted by the university last year, and so is not being considered for this section.

Nanyang performs relatively better with patents developed in-house than co-publications, and shows an especially high number of patents per 1000 students, while its performance is not as strong in terms of co-publication. Paris-Sorbonne tends to favour more solitary production of patents than in co-partnerships, perhaps reflecting the more explicit division of the public and private sectors in French society.

From this benchmark comparison, three distinct approaches to industrial relations emerge. From the US approach, which is very heavily geared towards the development of research to be passed over into the private sector for R&D, to the Chinese approach, which clearly favours industrial partnership in R&D, and more patents being developed in house, but in partnership with industry, and finally the more broadly European, and especially French, approach, which prefers to develop patents entirely in-house, using some partnerships for publication but not for sharing intellectual property in the form of patents.

USP, however, sits roughly in the middle of these different types; it favours co-patenting slightly more than co-publication, but not to an extent that a clear pattern emerges.

This type of analysis is not possible with ranking systems, as we cannot distinguish between ideal types if the metrics we are using are universal. There is no one approach which is the correct approach, this rather suggests that industrial relations are dictated by the politics of the nation state, internal

university governance structure, and the national systems of innovation into which the university finds itself inserted; i.e the extent to which the economy is structured to reward innovation in business. We might propose then, that it is more likely that universities in liberal market economies will be more likely to engage in co-publication than in patent development, simply because the economy rewards innovation better, and so it is more likely that firms will invest in their own R&D to turn the results of applied research produced with universities into patents of their own. For less sympathetic economies, it is more likely that firms will choose to develop the patent with the university, either due to a lack of skilled labour force and expertise, or as a way of hedging risk against expensive failures.

As we can see from the 'wheels' produced by U-Multirank, the sole Anglo-American entry on our shortlist, the University of California, scored significantly lower in international engagement and regional engagement, measured here as a percentage of publications as regional and international collaborations than the other seven listed institutions. This suggests that contrary to what we might expect, and what much of the literature suggests, it is in fact the regional powerhouses on the periphery that are more engaged in transnational collaborative enterprises than those on the centre.

We can, of course, point to the much larger pool of domestic researchers in the United States, it may be that US academics are less concerned with collaboration outside of their own environments, because their own environments are large enough to hold sufficient numbers of potential research collaborators. Compared to Nanyang, for example, which has only NUS, and perhaps one or two others as serious domestic peers, it is only natural that Singaporean researchers look more outside of their own country. Similarly, the Chinese C9 group of nine universities produces over 20% of its research output, and attracts over 10% of its total funding, despite representing just 2% of its researchers (YUE, ZHU; 2009). The benchmark suggests, despite the University of California's impressive research profile, it might not be the most relevant to compare to in terms of conducting transnational research, as other institutions are more intensively engaged in the formation of international partnerships.

At the time of writing, only universities within the European Union have begun to compare the teaching metrics, but a vague approximation of the metrics will have to suffice here. One of the huge problems facing Brazilian higher education as a whole, including USP, is the extremely high dropout rate of undergraduates. This has serious effects not just in resources wasted on individual students, but also on the function of the university as a teaching institution. There are a number of potentially explanatory internal and environmental factors, including financial hardship, a lack of properly structured curricula or improperly structured curricula, or a lack of proper pastoral or ongoing professional support. It has also been suggested that tuition fees have been an important factor in financially tying students to their own academic progress, leaving them more unwilling to abandon studies in which they have invested physically as well as intellectually. There are also a variety of cultural factors that can influence graduation rates, including whether deferral is a commonly accepted practice; in the UK it is only granted in extenuating circumstances and therefore dropout rates and deferral rates are much lower than in Brazil.

Graduation rates cannot be used as a metric to denote excellence as the number of students completing a course is not in any way indicative of the quality of the course that has been completed. It therefore could not be used as a metric in a ranking, however in benchmarking it could provide illustrative clues on an institutional level of how to reduce this inefficiency, and therefore increase the university's effectiveness as a teaching institution. We can see from the graduation rates in the table that USP lags far behind its competitors, with the exception of UNAM. All report graduation rates close to what would be expected (around 25% for a three or four year bachelors, around 33% for Copenhagen, in which a Bachelor's can be completed within two or three years, and so the percentage of the student body graduating each year would be expected to be a little higher).

While differences in graduation rates may be affected by the prevalence of mature students and students from lower socioeconomic backgrounds, who often require longer to complete courses for financial or personal reasons, it is clear that both USP and UNAM's impaired ability to produce graduates within

the expected (normative) time period means that the importance of student spend is lessened if students take much longer to complete their degrees. Large scale recording of this information, including individual institutions managing significant improvements in student adherence would allow much more widespread exchanges of successful policies of student support and teaching programme structure.

For this benchmark, employment rates will not be considered as they are in the U-multirank for the simple reason that the tool was developed for use primarily within the EC and institutions inserted into the Bologna process. This means that, with some variations, most institutions participate in a common labour market with relatively similar economic conditions. Although graduates' readiness for the workforce is something that could be analysed in the context of a benchmarking exercise, the comparison would be difficult here, also owing to the lack of published or recorded information on graduate employment rates.

The research metrics on citation rates and publication numbers for U-Multirank are not being considered in any profound way, because of the problems with using citation rate as a proxy for performance on a universal scale. However, both USP and UNAM demonstrate a very high number of papers published both in absolute numbers, and in terms of productivity, but lag far behind in terms of translating this to bibliometric performance.

When comparing within systems, for example comparing American institutions with other American institutions citation rates may still be indicative of individual achievement, but for the developing world, or the educational periphery, they are artefacts of a university's position in the network, not reflections of their individual contributions.

APPENDIX B: MAPPING CLIMATE CHANGE

Climate change is perhaps the single greatest challenge to future prosperity, both human and biological. Universities can, and must, respond to the challenge of producing research and solutions to it. It has been chosen as a research area for this analysis because of this, but furthermore because it fits the model for both Mode 2 and new sciences (BONNACORSI, 2001). Research being carried out on climate change spans many disciplinary areas, from biological and earth sciences to human and social sciences, over a broad array of topics demanding a different type of measurement and integration of efforts than previous challenges. One of the problems of the evaluation methods outlined above is that it cannot properly account for or stimulate this type of work, however using Altmetrics and social network mapping it is possible to eliminate disciplinary bias and provide much more descriptively rich analytics of a complex problem (BORNMANN, 2014; HOFFMANN; LUTZ; MECKEL, 2014).

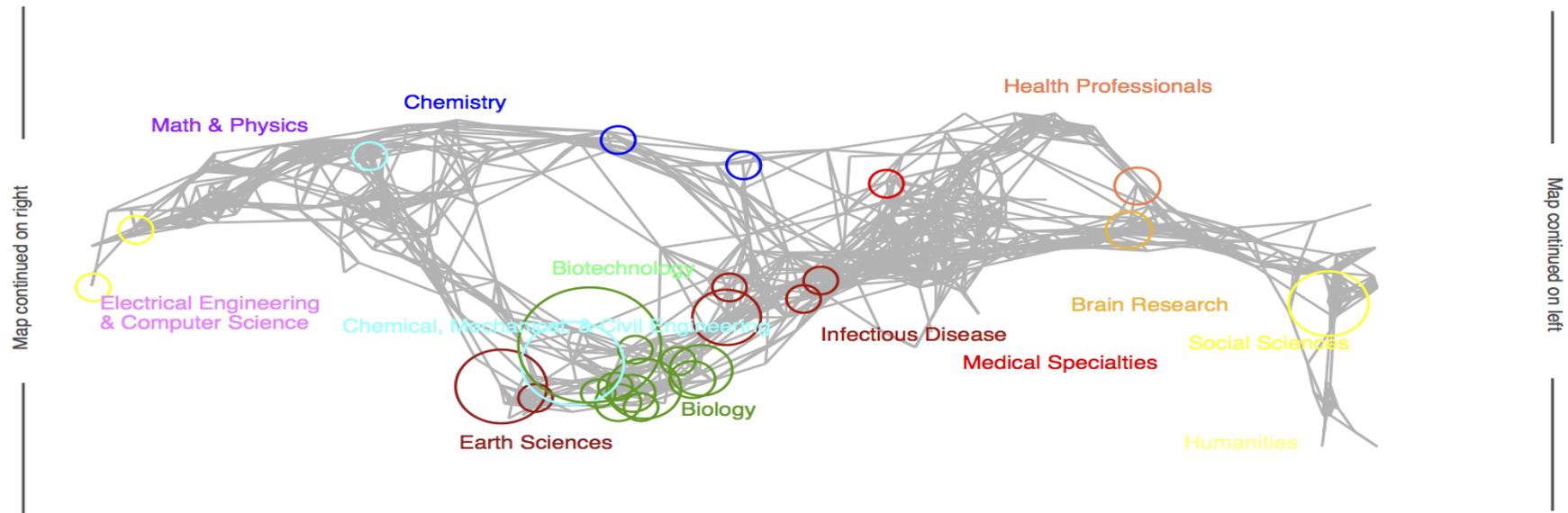
A search was performed on the Web of Science database using the search terms “ADDRESS: Universidade de São Paulo” and “TOPIC: climate change” and “TOPIC: mudança climática”, which produced 170 unique results. These were then processed and consolidated in the program Sci2 (Sci2 Team. (2009). Science of Science (Sci2) Tool. Indiana University and SciTech Strategies, <https://sci2.cns.iu.edu>), and their results mapped onto the UCSD Map of Science (available from <http://sci.cns.iu.edu/ucsdmap/>). This gives us a substantial idea of what USP's approach to climate change is in terms of its published output.

Much of USP's research was in uncategorised journals, mainly because the SciELO network is still a relatively recent entrant to the WoS network, and so has not received the same amount of attention as the WoS core collection. This gives the appearance of a much more starkly populated map than the other two, who publish in more mainstream sources. In fact, Nanyang returned 100 results, and University of California returned 266, compared to USP's 170. This is a limitation on the analysis, but not necessarily a serious one, because this is a representation of active areas of research, not production.

Fig. 1: USP “Climate Change” output on Web of Science (2015)

Topical Visualization

Generated from Standardized Journal Names
 40 out of 77 records were mapped to 27 subdisciplines and 9 disciplines.
 April 26, 2015 | 02:52 PM BRT

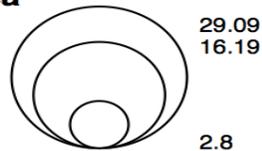


□ 2008 The Regents of the University of California and SciTech Strategies.
 Map updated by SciTech Strategies, OST, and CNS in 2011.

Legend

Circle area: Fractional record count
 Unclassified = 37
 Minimum = 1
 Maximum = 24
 Color: Discipline
 See end of PDF for color legend.

Area



How To Read This Map

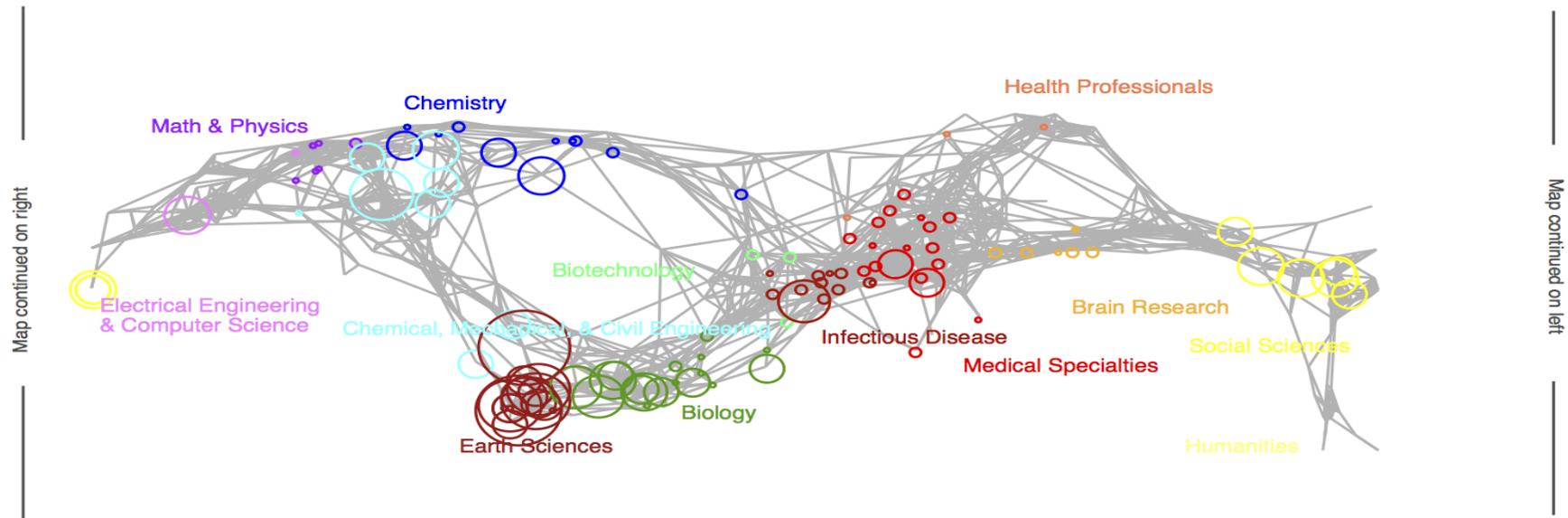
The *UCSD map of science* depicts a network of 554 subdiscipline nodes that are aggregated to 13 main disciplines of science. Each discipline has a distinct color and is labeled. Overlaid are circles, each representing all records per unique subdiscipline. Circle area is proportional to the number of fractionally assigned records. Minimum and maximum data values are given in the legend.

CNS (cns.iu.edu)

Fig 2. Nanyang Technological University “Climate Change” output on Web of Science (2015)

Topical Visualization

Generated from Standardized Journal Names
 73 out of 80 records were mapped to 112 subdisciplines and 12 disciplines.
 April 26, 2015 | 02:41 PM BRT

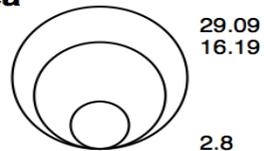


© 2008 The Regents of the University of California and SciTech Strategies.
 Map updated by SciTech Strategies, OST, and CNS in 2011.

Legend

Circle area: Fractional record count
 Unclassified = 7
 Minimum = 0
 Maximum = 9
 Color: Discipline
 See end of PDF for color legend.

Area



How To Read This Map

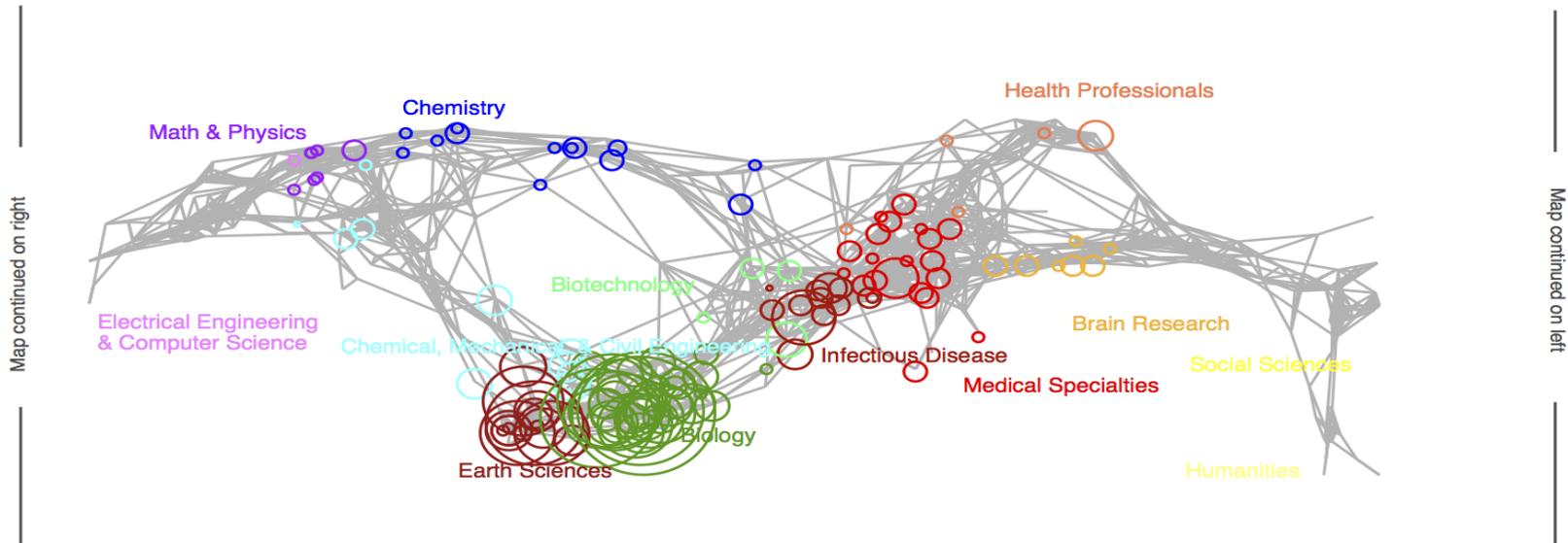
The *UCSD map of science* depicts a network of 554 subdiscipline nodes that are aggregated to 13 main disciplines of science. Each discipline has a distinct color and is labeled. Overlaid are circles, each representing all records per unique subdiscipline. Circle area is proportional to the number of fractionally assigned records. Minimum and maximum data values are given in the legend.

CNS (cns.iu.edu)

Fig. 3 University of California “Climate Change” output on Web of Science (2015)

Topical Visualization

Generated from Standardized Journal Names
 118 out of 153 records were mapped to 134 subdisciplines and 11 disciplines.
 April 26, 2015 | 03:06 PM BRT

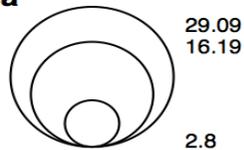


© 2008 The Regents of the University of California and SciTech Strategies.
 Map updated by SciTech Strategies, OST, and CNS in 2011.

Legend

Circle area: Fractional record count
 Unclassified = 35
 Minimum = 0
 Maximum = 24
 Color: Discipline
 See end of PDF for color legend.

Area



How To Read This Map

The *UCSD map of science* depicts a network of 554 subdiscipline nodes that are aggregated to 13 main disciplines of science. Each discipline has a distinct color and is labeled. Overlaid are circles, each representing all records per unique subdiscipline. Circle area is proportional to the number of fractionally assigned records. Minimum and maximum data values are given in the legend.

CNS (cns.iu.edu)

When placed next to the University of California and Nanyang maps, one thing becomes immediately clear. Although all three have an emphasis mainly on biology and earth sciences, USP has a pronounced profile in social sciences, especially in international politics and agrarian sciences. Nanyang has a focus on both this, and economics. This is entirely absent from the UC map, showing that for the UC research program, climate change is principally a technological and scientific problem that has yet to be linked to the human or social dimensions of research, as populations in the US are less vulnerable than in the developing world, the issues aren't being specifically linked. Because of this, we can see that perhaps south-south cooperation would be more beneficial for areas of social benefit, while the high concentration of work published in infectious disease treatment at UC would be an area worth exploring with UC.

These maps show the danger of using citation count as a measure of these more transdisciplinary areas, as the citation rates and behaviour are significantly different between areas, and so while two articles may be ostensibly on the same topic, climate change, we should expect significant differences in publishing and citation behaviour between a Brazilian social sciences article and a bio-technology article published in the US.

8.2.2 Keyword frequency and Linkage

Keyword data are difficult to manipulate due to the fact that many researchers use tiny variations in similar themes for their papers. If keywords could be systematised in some way, analysis of central themes within schools could be measured with much greater granularity, we could understand linkages between issues and schools. For publication on the Web of Science, Thomson Reuters keeps a list of over 54,000 permitted keywords to tag to articles, ensuring that this data can be analysed in a much more structured way. However, for the purposes of this exercise, the software Sci2 and NWB were used to data mine and standardise the abstracts, before processing them with the DrL (VxORD) (MARTIN et al. 2001) algorithm to produce an undirected

just one single view of climate change research at USP, but a variety of perspectives that allow for greater understanding of dynamics at work, and a more nuanced understanding of how impact can be represented. Co-citation networks like this one were used in a project by the Ministério de Saúde as a novel means of evaluating and identifying capacities for research of neglected diseases such as the aforementioned schistosomiasis to aid identification of key research actors, centres and vital cutoff points, recognising that traditional means of assessment were of limited use in identifying competences in areas of low publication. These maps are not complex to create once the initial algorithm is established, and can offer much greater insight into the way in which scientific communities interact with one another, allowing for more sophisticated research steering and evaluation (MOREL, SERRUYA, PENNA, GUIMARAES; 2009). The benefits of working in this way was noted by the researchers, who were more able to effectively motivate key stakeholders and orient research committees, with a resultant upturn in production.

8.2.4 Interdisciplinarity

There are many experimental measures of interdisciplinarity that have been trialled over recent years (LAWSON; 2013), and it should be noted that the majority are still in their experimental, and nascent phases, and so wide ranging conclusions are not always easy to draw. Furthermore, at present no one methodology can measure the profundity of the disciplinary links as well; whether are truly interdisciplinary or analyzing issues across disciplinary boundaries but still from a broadly disciplinary perspective, or more importantly, whether the research attained its goals, if we accept that one of the main drivers of transdisciplinary research is that it is applied to context of its problem. There have been attempts to resolve these problems through the introduction of qualitative survey assessments like the T-competence test, but these are of limited use to large scale surveying.

Perhaps the most complete attempts yet to represent multidisciplinary have been the overlay maps of the Web of Science index (RAFOLS; 2012). These are the template for the climate change maps presented above, with the added benefit of being able to measure not just the publication location, which is a rather blunt measure of diversity in approaches to a topic, he can also measure flows of citation between different disciplinary journals, both as inputs (variety and density of inter-disciplinary citation in the work) and as outputs (variety and density of inter-disciplinary citation that the work receives).

Rao-Stirling Diversity (STIRLING: 2002) is a score of diversity measured by diversity of publication areas, balance between production in different areas and cognitive distance between areas. This is widely regarded as the most reliable single indicator of interdisciplinarity.

8.2.5 Geospatial Mapping Analyses

While it is commonplace to bemoan the geographical dynamics at work in Latin America, which gives us a natural disadvantage over, say, the United States, little attempt up to now has been made to identify and explain these dynamics. Particularly in Europe, great strides have been made in recent years

to explain these dynamics (FRENKEL, HOEKMAN; 2013) by mapping collaborations, production locations and citation flows onto geographical maps in projects such as Living Science (livingscience.inn.ac), and excellence heat maps for specific fields. Thinking spatially about citation impact and publication location would allow us to visually represent tendencies, produce expected impact values in a given field, better understand the impacts of dynamics like academic mobility (ZUBIETA; 2012) and international collaboration. For climate change, a comprehensive quantitative study (PASGAARD, STRANGE; 2013) showed that research is disproportionately carried out by rich countries without vulnerable populations. This means that Brazil provides a crucial perspective and voice to climate change research, but also that it generally hasn't received as much attention as research produced in the North.