

Avaliação Acadêmica Multidimensional com o uso do “U-Multirank”

Multidimensional Academic Evaluation using the "U-Multirank"

Antonio Fernando Bertachini de Almeida Prado¹

Resumo: A avaliação de instituições Acadêmicas é um tópico de grande interesse no mundo e sempre trouxe muitas polêmicas. Esse assunto é atual no Brasil, pois a CAPES estuda novas formas de avaliação para a Pós-Graduação brasileira. A ideia principal é usar um modelo multidimensional, de forma a valorizar aspectos que são importantes para instituições Acadêmicas e para as comunidades onde estão inseridas, mas que não são captados pela avaliação atual, bastante focada em resultados gerados por pesquisas de ponta. Na ausência de uso de métodos multidimensionais de avaliação acadêmica no Brasil, é interessante analisar experiências internacionais nessa área para aprimorar o entendimento de como eles funcionam. Nesse cenário destaca-se o ranking acadêmico internacional “U-Multirank”, criado na Europa e que tem resultados desde 2014. O presente estudo analisará as Dimensões e Indicadores desse ranking e mostrará as correlações existentes entre Dimensões e Indicadores nas instituições de Ensino Espanholas.

Palavras-chave: rankings acadêmicos; avaliação da educação superior; avaliação multidimensional; gestão universitária.

Abstract: Evaluation of Academic Institutions is of great interest in the world, always causing controversies. This subject is current in Brazil, because CAPES is studying new forms to evaluate Brazilian Graduate Programs. The main idea is to use multidimensional models, to consider aspects that are important for academic institutions and for the communities where they are inserted, but that are not captured by the current evaluation, which is focused on results generated by international research. In the absence of multidimensional evaluation methods in Brazil, it is interesting to analyze international experiences in this area, to improve our understanding on how they work. In this scenario, the international academic ranking “U-Multirank” stands out. It was created in Europe and has results available since 2014. The present paper will analyze the Dimensions and Indicators of this ranking and their correlations in Spanish Education Institutions.

Keywords: academic rankings; higher education evaluation; multidimensional evaluation; university management.

¹ Instituto Nacional De Pesquisas Espaciais| São José dos Campos| SP| Brasil.
Contato:antonio.prado@inpe.br. ORCID: <https://orcid.org/0000-0002-7966-3231>

- Recebido em: 28 de julho de 2021
- Aprovado em: 21 de outubro de 2021

DOI: <http://dx.doi.org/10.1590/S1414-40772022000100009>

Este é um artigo publicado em acesso aberto sob uma licença Creative Commons
<https://creativecommons.org/licenses/by-nc/4.0/>

Introdução

Avaliação de instituições Acadêmicas é sempre um tópico importante em educação e política científica, e tem sido alvo de grandes debates. Instituições acadêmicas têm missões, valores e objetivos bastante diferentes entre si, e a tarefa de criar Indicadores únicos para uma avaliação geral de todas elas é bastante complexa.

Por essa razão, a avaliação de programas de pós-graduação, feita regularmente pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), sempre foi alvo de muitas críticas da comunidade científica pelo seu enfoque excessivamente voltado para a pesquisa e internacionalização. Para trabalhar na solução desses problemas, tem sido estudada, nos últimos anos, uma mudança de critérios de avaliação, com a inclusão de novas dimensões que possam valorizar outros aspectos além de pesquisa. O objetivo é ter uma avaliação mais qualitativa e que considere Indicadores de impacto social, usando as componentes quantitativas apenas como um guia geral e não como um elemento decisivo na atribuição de notas. Essas novas regras terão uma grande influência nas notas finais dos Programas de Pós-Graduação do Brasil, impactando diretamente na sobrevivência dos mesmos a médio e longo prazo. Essas alterações serão sentidas por Programas de Pós-Graduação de todas as áreas, mas serão particularmente fortes nos Programas das áreas de ciências, cuja maioria está estruturada em termos de produção científica em periódicos internacionais de alto impacto, principalmente os Programas mais bem avaliados no momento.

Para fazer essa mudança, as novas regras devem ser baseadas em uma avaliação multidimensional, e faz-se necessário o estudo de Dimensões e Indicadores que sejam adequados a esse fim. Como esse tipo de avaliação tem pouca tradição no Brasil, estudar opções em uso no exterior é de grande importância no momento. Um destaque nesse tipo de avaliação é o ranking multidimensional europeu “U-Multirank”. Ele usa cinco Dimensões divididas em 36 Indicadores e avalia instituições Acadêmicas do mundo todo, embora com maior participação de países europeus. A CAPES tem observado esse ranking (CAPES, 2019), e vários dos Indicadores utilizados por ele estão sendo propostos para a nova forma de avaliação da Pós-Graduação brasileira. Alguns desses Indicadores estão incluídos na avaliação referente ao período 2017-2020, uma avaliação considerada como de transição.

Para ajudar nesse entendimento, o presente trabalho faz uma revisão das Dimensões e Indicadores usados por esse ranking, mostrando suas formas de obtenção, caráter objetivo ou subjetivo, problemas ligados a falta de dados, etc. Para efetuar um estudo mais detalhado, a Espanha foi escolhida como amostra, por ser um país com um elevado número de instituições

Acadêmicas listadas nesse ranking (79), ser um país europeu, onde o “U-Multirank” foi criado e é mais conhecido e valorizado, logo tem uma base de dados mais completa, e também por ter similaridades culturais com o Brasil. Os resultados vão mostrar como se comportam os índices de correlação entre Dimensões para essas instituições, dividindo-as em faixas conforme seu desempenho global. Isso significa verificar, para cada faixa de desempenho estudada, se as instituições possuem um nível maior de especialização nas suas Dimensões mais fortes ou se tem um desempenho homogêneo entre os diferentes aspectos avaliados.

Após esse estudo do “U-Multirank”, serão feitas algumas considerações gerais sobre o possível uso dessa metodologia pela CAPES, comentando as vantagens e desvantagens de cada Indicador. O trabalho não está focado em construir um novo modelo de avaliação para a Pós-Graduação brasileira, tarefa bastante complexa, mas serão apontados alguns pontos positivos e negativos sobre cada Dimensão e Indicador, do ponto de vista de aplicação na Pós-Graduação brasileira, visando colaborar com o processo de adaptação desse método.

Histórico da avaliação da Pós-Graduação brasileira

A CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) conduz uma avaliação dos Programas de Pós-Graduação no Brasil desde 1976 (NOBRE; FREITAS, 2017). Inicialmente, as avaliações eram baseadas em visitas aos programas e seus resultados mantidos em sigilo, com cada programa tendo conhecimento apenas da sua nota, como meio de buscar uma melhoria de sua qualidade. A partir do final dos anos oitenta, os resultados passaram a ser divulgados publicamente, gerando com isso muita polêmica sobre os métodos utilizados nas avaliações (CASTRO; SOARES, 1983).

Outra mudança significativa no processo avaliativo da pós-graduação ocorreu em 1998, na avaliação referente ao biênio 1996-1997, quando critérios mais exigentes passaram a vigorar. Com isso, muitos programas tradicionalmente bem avaliados tiveram quedas bruscas em suas notas. A avaliação passou a ser bastante centrada em “pesquisa” e “internacionalização”. O conceito de “Qualis”, que é uma classificação de periódicos que visa avaliar a importância relativa de cada periódico para fins de pontuação na avaliação, passou a ter um peso muito significativo na nota final. Em essência, a pós-graduação brasileira passou a ser avaliada com um foco muito grande nos aspectos “pesquisa” e “internacionalização”, entendendo que o lado “ensino” é o meio pelo qual se produz pesquisa de alto nível, e não um fim em si mesmo. Esse enfoque gerou muitos debates e críticas de muitas áreas (FONSECA; OLIVEIRA, 2010; FURTADO; HOSTINS, 2014; KAWASAKI, 2017; PATRUS; SHIGAKI; DANTAS, 2018). O processo avaliativo teve alterações ao longo dos anos, como resposta a

muitas dessas críticas, mas sua essência não teve alteração substancial até a avaliação efetuada para o quadriênio 2013-2016.

Destaca-se que a avaliação da CAPES tem grande importância na definição de políticas educacionais brasileiras, pois dela dependem a validação dos diplomas emitidos pelas instituições brasileiras, o número de bolsas de mestrado, doutorado e pós-doutorado concedidas aos programas pela CAPES, assim como muitos outros aportes de recursos feitos aos Programas de Pós-Graduação. Além disso, Programas de Pós-Graduação com doutorado necessitam uma nota mínima 4 para terem seus diplomas reconhecidos pela CAPES, enquanto que Programas que possuem apenas o mestrado necessitam de nota mínima 3. Outras agências de fomento também usam a nota da CAPES como critério na distribuição de recursos. Esses fatos explicam a grande importância de uma discussão mais detalhada dos métodos de avaliação brasileiros e internacionais, mesmo que todos eles sofram críticas das comunidades científica e educacional.

Além desses fortes fatores econômicos, melhores notas se transformam em atrativos para candidatos a alunos. Isso faz com que os Programas com as melhores notas atraiam os melhores alunos, e isso cria um “círculo virtuoso” para Programas bem avaliados, enquanto que Programas com piores avaliações entram em um “círculo de queda”. Nesse sentido, a avaliação efetuada pela CAPES se torna um elemento muito forte como diretriz de política educacional. A ficha de Indicadores elaborados pela área avaliativa da CAPES acaba sendo um modelo a ser seguido por todos os cursos, sob pena de fechamento. Essa ficha define um padrão para o número e duração de teses e dissertações, quantidade de linhas de pesquisa e projetos, relação docentes/discentes, média de publicações por docentes e discentes, etc. Esse padrão passa a atingir também os docentes, pois para que possam permanecer credenciados em Programas de Pós-Graduação eles precisam atingir critérios individuais de produtividade.

Histórico de rankings internacionais

Os rankings internacionais de avaliação acadêmica surgiram na década de 2000 (CALDERÓN; FRANÇA, 2018), e tiveram como objetivo identificar instituições Acadêmicas que pudessem ser consideradas como de “classe mundial” (MOURA; MOURA, 2013; LLOYD; ORDORIKI; RODRIGUES-GÓMEZ, 2011; SHIN; TOUTKOUSHIAN; TEICHER, 2011; STACK, 2016; BARREYRO, 2008; PILATTI; CECHIN, 2018; MENDES; DUTRA, 2019; SOUZA, 2017). O primeiro ranking desse tipo a surgir foi o “Academic Ranking of World Universities - (ARWU)” (<http://www.shanghairanking.com/ARWU2020.html>), também conhecido como “Ranking de Xangai”, criado em 2003 pela Universidade de

Xangai, na China (CALDERÓN; FRANÇA, 2018; THERY, 2010). Ele foi criado com o objetivo de fornecer informações ao governo da China para ajudar no processo de escolha de instituições de Ensino internacionais para o envio de estudantes chineses ao exterior e também para medir a qualidade das instituições chinesas. A China já tinha um programa muito forte de envio de estudantes ao exterior nessa época, visando melhorar a qualidade de suas instituições de ensino e pesquisa nos anos seguintes, após o retorno desses estudantes para a China.

Inspirado nesse trabalho, que ganhou grande visibilidade mundial, surgiram outros rankings internacionais, tais como o “Webometrics Ranking of World Universities” (<http://www.webometrics.info/en>) em 2004 e o THE-QS, também em 2004, que seria subdividido em THE (<https://www.timeshighereducation.com/world-university-rankings>) e QS (<https://www.topuniversities.com/university-rankings>) em 2010. O Brasil também criou alguns rankings acadêmicos como, por exemplo, o ranking elaborado pelo jornal Folha de São Paulo (RIGHETTI, 2015). Uma descrição histórica sobre rankings acadêmicos internacionais pode ser encontrada em Andriola e Araújo (2018).

Esses rankings recebem muitas críticas de especialistas em educação e a literatura tem muitas polêmicas sobre sua aplicação e uso. Calderón e França (2018) mostra um estudo detalhado da literatura e dos argumentos contra e a favor de rankings acadêmicos. Porém, com a globalização da economia e da educação no mundo, eles ganharam muita popularidade e atualmente são inevitáveis. Eles fazem parte da busca por internacionalização da educação, um fenômeno mundial. Cada vez mais os resultados obtidos nesses rankings ganham grande destaque na mídia (DRESCH, 2018; ROTHEN; SANTANA, 2018). Quase todas as instituições de Ensino e pesquisa no Brasil e no mundo fazem esforços para melhorar suas classificações nesses rankings e dão grande destaque em seus veículos de comunicação quando obtém bons resultados. Passou a existir uma grande competição para ter reconhecimento da população em geral, tanto por parte de instituições públicas como privadas.

Existem outros eventos importantes que mostram a pressão que instituições de Ensino e professores brasileiros sofrem por internacionalização. Um bom exemplo é o programa de Internacionalização lançado pela CAPES em 2017 (PrInt) (CAPES, 2017), que concedeu recursos para mobilidade internacional e custeio para 36 instituições de Ensino e pesquisa no Brasil. Esse projeto deu grande visibilidade as instituições que foram contempladas, além dos recursos financeiros.

Introdução ao “U-Multirank”

O “U-Multirank” europeu (VAN VUGHT; ZIEGELE, 2012; PRADO, 2021a, 2021b, 2021c), tem como um dos grandes atrativos o seu caráter multidimensional. Isso permite uma avaliação em um leque mais amplo, olhando para um número maior de Indicadores. Isso faz com que estudos preliminares desse ranking sejam importantes, no sentido de aumentar nosso conhecimento sobre ele. Destaca-se aqui a forte ligação que passa a existir entre rankings internacionais e a avaliação da pós-graduação brasileira, já que ambos buscam Dimensões e Indicadores para medir desempenho acadêmico.

O “U-Multirank” não tem como objetivo elaborar uma classificação geral das instituições de Ensino. Não define a melhor instituição Acadêmica do mundo, ou de um dado país. As classificações são feitas somente em cada um dos Indicadores específicos de desempenho, com muitas instituições tendo notas iguais, pelo fato das notas serem atribuídas em apenas cinco valores, como será detalhado mais a frente. O ranking tem como objetivo apresentar desempenhos institucionais mostrando pontos fortes e fracos de cada uma, identificando perfis específicos, sem definir melhores e piores. A filosofia de trabalho é baseada na hipótese de que não existem “melhores” e “piores” instituições de Ensino, mas sim as “mais” e “menos” adequadas ao interesse de cada aluno. Com isso, cada usuário pode fazer seu próprio ranking, utilizando os Indicadores que considerar mais importantes. Outra justificativa apresentada para a multidimensionalidade é que os rankings unidimensionais não são robustos, dado que pequenas modificações nos pesos dos Indicadores atualmente utilizados alteram significativamente os resultados, o que reduz muito a validade dessas classificações unidimensionais.

No “U-Multirank” as instituições são classificadas em cinco grupos de desempenho em cada Indicador (A, B, C, D, E), com A expressando melhor desempenho. Isso é feito para ajudar a eliminar problemas de precisão de medida, já que é feito um agrupamento de instituições em bandas mais largas. Com isso, o problema de precisão aparece apenas nas linhas divisórias entre as faixas de notas, sendo assim minimizado.

Os dados utilizados pelo sistema são extraídos de informações fornecidas pelas próprias instituições (questionários), bancos de dados bibliométricos e de patentes internacionais, bancos de dados nacionais e pesquisas respondidas por aproximadamente 145.000 estudantes (questionários). A análise bibliométrica é baseada no banco de dados Web of Science (WoS) (Science Citation Index Expanded, Social Sciences Citation Index e Arts & Humanities Citation Index). O ranking possui mais de 1.700 instituições de Ensino de 92 países na versão de 2020.

Uma análise dos Indicadores mostra que alguns deles são obtidos de fontes abertas, tais como número de publicações, citações, etc. Essas informações se encontram disponíveis para todas as instituições de Ensino analisadas e tem alto grau de confiabilidade. Porém, informações tais como local de trabalho de egressos e tempo de titulação, são obtidas de questionários enviados pelas instituições de Ensino e estudantes. Esses dados nem sempre estão disponíveis e não tem um alto grau de confiabilidade e precisão.

Para mostrar as particularidades do “U-Multirank”, foi feito um comparativo entre esse ranking e os mais tradicionais QS e THE (PRADO; AGUILAR, 2021). Destaca-se desse trabalho que o “U-Multirank” tem 5 Dimensões, enquanto que o QS e o THE são unidimensionais. Do ponto de vista de Indicadores utilizados, o QS utiliza 6, o THE 13 e o “U-Multirank” 36, o que gera uma visão muito mais abrangente do desempenho das instituições de Ensino. Ele inclui instituições de menor porte com forte atuação regional e local, o que ajuda o trabalho de busca para futuros alunos. Do ponto de vista de disponibilidade de dados, THE e QS possuem 100% dos dados disponíveis, enquanto que o “U-Multirank” teve apenas 54.46% em 2020. Outra característica do “U-Multirank” é que ele não tem nenhum dado subjetivo, baseado em “reputação acadêmica”. O QS tem 50% de sua nota baseado em “reputação acadêmica”, enquanto que o THE tem 33%. Quanto à forma de obtenção dos dados, os três rankings enviam questionários às instituições de ensino e usam bases de dados abertas, tais como Scopus e Web of Sciences. O THE e o QS enviam questionários à comunidade acadêmica internacional para medir “reputação acadêmica”. Apenas o THE envia questionários a empregadores, perguntando quais instituições de ensino formam alunos inovadores e com boa formação acadêmica. Apenas o “U-Multitank” envia questionários a alunos e ex-alunos, como forma de verificar as informações enviadas pelas instituições de Ensino.

Medindo a quantidade de dados disponíveis do “U-Multirank”

Para realizar um estudo mais geral sobre a quantidade de dados disponíveis e verificar os pontos fracos devido à ausência de dados vamos fazer uma análise geral sobre esse tópico em todas as Dimensões e Indicadores do “U-Multirank”. Vamos considerar todos os dados disponíveis na versão de 2020 do ranking abrangendo todos os países disponíveis.

A Tabela 1 mostra todos os Indicadores utilizados pelo “U-Multirank” divididos nas cinco Dimensões. Ela mostra a descrição da Dimensão, as quantidades de dados disponíveis, não disponíveis e que não se aplicam, bem como a fonte de obtenção dos dados (questionários respondidos pelas instituições ou dados abertos).

Algumas observações podem ser feitas. Em primeiro lugar destaca-se o grande número de informações não disponíveis. Considerando todas as 1786 instituições Acadêmicas listadas no ranking de 2020 e considerando a existência de 36 Indicadores, deveríamos ter 64.296 notas. Porém, nota-se que temos apenas 35.016, o que corresponde a 54,46%. Temos 25.618 dados faltantes (39,84%) e 3.662 dados (5,70%) que foram considerados como “não aplicáveis”.

Além de volumosa, a distribuição de ausências dos dados não é uniforme, fazendo com que alguns Indicadores sejam mais completos que outros. A Tabela 1 mostra em detalhes a situação de cada Indicador e a Figura 1 mostra o número de dados disponíveis por Indicador. Uma leitura dessas informações mostra um grande desbalanceamento entre os Indicadores para os dados disponíveis. Nota-se uma clara lacuna entre os valores de 1.000 e 1.400 dados disponíveis, de um máximo de 1.786 dados por Indicador. Sendo assim, é possível traçar uma linha em torno do valor 1.428, que corresponde a 80% dos dados, e dividir os Indicadores entre os que estão acima ou abaixo dessa linha. Na Tabela 1 esses Indicadores estão marcados em negrito. Assumindo esses Indicadores como melhores para uma avaliação comparativa de todas as instituições listadas, não teríamos nenhum Indicador adequado para a Dimensão **Ensino e Aprendizagem** (de um total de 4 Indicadores), 6 Indicadores adequados para a Dimensão **Pesquisa** (de um total de 11 Indicadores), 4 Indicadores adequados para a Dimensão **Transferência de conhecimento** (de um total de 9 Indicadores), 1 Indicador adequado para a Dimensão **Internacionalização** (de um total de 6 Indicadores) e 2 Indicadores adequados para a Dimensão **Engajamento Regional** (de um total de 6 Indicadores).

As principais razões dessa falta de dados são o grande número de Indicadores utilizados pelo “U-Multirank” e o fato de que uma parcela significativa desses dados é obtida através de respostas a questionários enviados para as instituições de ensino, muitas das quais não enviam respostas ou enviam apenas parcialmente. Isso não é um problema grave para o ranking, pois ele não tem por objetivo elaborar uma classificação geral, mas sim ser um guia para futuros alunos escolherem suas Universidades. Nesse sentido, o futuro aluno pode simplesmente não considerar os dados faltantes e fazer sua decisão com base nos dados disponíveis. Existe uma perda de informação, mas ela não inviabiliza o uso do ranking.

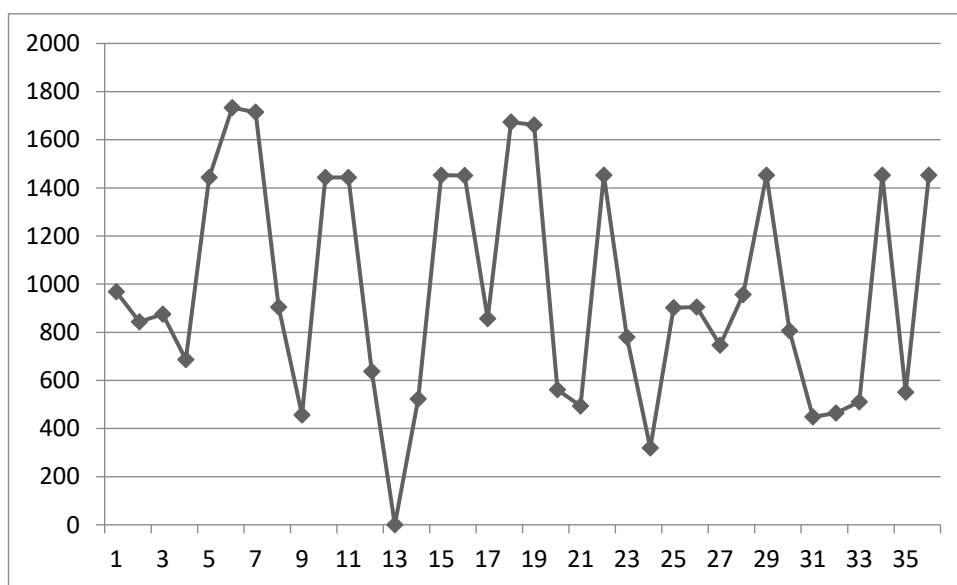
A Tabela 1 também mostra, em itálico, Indicadores que não aparecem na forma gráfica clássica de apresentação dos resultados do “U-Multirank”, mas que são apresentados nas tabelas comparativas efetuadas pelo “U-Multirank”.

Tabela 1 – Indicadores do “U-Multirank” mostrando os dados disponíveis, os dados que não se aplicam e a fonte dos dados: QI indica questionário respondido pela instituição de Ensino e AB indica dado aberto (acesso público). Negrito destaca os dados com mais de 80% de disponibilidade e itálico mostra os dados que não aparecem na forma gráfica

Ensino e Aprendizagem	Dados disponíveis	Dados não disponíveis	Dados que não se aplicam	Fonte dos dados
Taxa de graduação no bacharelado	968	818	0	QI
Taxa de graduação no mestrado	843	943	0	QI
Graduação no tempo esperado (bacharéis)	875	911	0	QI
Graduação no tempo esperado (mestrado)	686	1100	0	QI
Pesquisa				
Taxa de citação	1443	62	281	AB
<i>Publicações (números absolutos)</i>	1733	53	0	AB
Publicações (tamanho normalizado)	1714	72	0	AB
Receita de pesquisa externa	905	881	0	QI
Produtos relacionados à arte	456	1329	1	QI
Publicações mais citadas	1443	62	281	AB
Publicações interdisciplinares	1443	62	281	AB
Vagas de pós-doutorado	637	1147	2	QI
<i>Parcerias de pesquisa estratégica</i>	<i>0</i>	<i>1786</i>	<i>0</i>	<i>QI</i>
<i>Publicações Profissionais</i>	<i>523</i>	<i>1263</i>	<i>0</i>	<i>QI</i>
<i>Publicações de acesso aberto</i>	1452	53	281	AB
Transferência de conhecimento				
Publicações em co-autoria com parceiros industriais	1451	54	281	AB
Renda de fontes privadas	857	929	0	QI
<i>Patentes concedidas (números absolutos)</i>	1674	112	0	AB
Patentes concedidas (tamanho normalizado)	1662	123	1	AB
Patentes em co-autoria com a indústria	561	112	1113	AB
Spin-offs	493	1293	0	QI
Publicações citadas em patentes	1452	53	281	AB
Rendas obtidas de desenvolvimento profissional contínuo	779	1007	0	QI
<i>Empresas criadas por formandos</i>	<i>320</i>	<i>1466</i>	<i>0</i>	<i>QI</i>
Internacionalização				
Programas de bacharelado em língua estrangeira	902	876	8	QI
Programas de mestrado em língua estrangeira	905	879	2	QI
Mobilidade estudantil	746	1040	0	QI
Equipe acadêmica internacional	957	829	0	QI
Publicações conjuntas internacionais	1452	53	281	AB
Porcentagem de Doutores Estrangeiros formados	806	974	6	QI

Engajamento Regional				
Formados na Graduação trabalhando na região	448	1338	0	QI
Formados no Mestrado trabalhando na região	464	1322	0	QI
Estágios de estudantes na região	511	1275	0	QI
Publicações conjuntas regionais	1452	53	281	AB
Renda de fontes regionais	551	1235	0	QI
<i>Publicações regionais com parceiros industriais</i>	1452	53	281	AB

Figura 1 – Dados disponíveis por Indicador do “U-Multirank”



A seguir, a Tabela 2 mostra a quantidade de dados abertos e de dados fornecidos pelas instituições de Ensino. Observamos que existem 22 Indicadores obtidos através de questionários respondidos pelas instituições e 14 Indicadores obtidos de dados abertos. Isso indica uma forte dependência das notas e das classificações com relação aos dados informados pelas instituições de ensino. Isso significa que uma base completa de dados é fundamental, e a aplicabilidade do “U-Multirank” varia entre as regiões geográficas, pois a falta de dados da base varia muito de um país para outro.

Tabela 2 - Quantidade de dados abertos (AB) e fornecidos pelas instituições (QI) por Dimensão do “U-Multirank”

Dimensão	D1	D2	D3	D4	D5
QI	4	5	4	5	4
AB	0	6	5	1	2

Correlações entre os desempenhos das instituições de Ensino da Espanha para todas as Dimensões do “U-Multirank”

O próximo ponto a ser estudado é o comportamento das correlações estatísticas entre as diferentes Dimensões do “U-multirank”, como uma forma de prever seu comportamento e entender se as instituições estudadas têm tendência a terem comportamentos homogêneos ou heterogêneos com relação aos desempenhos entre as diferentes Dimensões.

Foram utilizadas as 79 instituições de Ensino da Espanha avaliadas pelo “U-multirank” na edição de 2020. A escolha dessa amostra é feita pelo fato da Espanha ter um número grande de instituições listadas e a base de dados tem uma porcentagem de dados “não disponíveis” bem abaixo da média geral. Dos 2.528 conceitos esperados, 2.187 estão disponíveis, o que significa 86,15% de dados completos, contra 54,46% para o total de instituições listadas no ranking de 2020. Pesa também o fato de, ao escolher instituições de um mesmo país, pode-se eliminar boa parte dos efeitos devidos a diferentes culturas, idiomas, legislações nacionais e outras particularidades, deixando o foco principal nos Indicadores e Dimensões. Com essa classificação, dividimos as instituições de ensino em quatro faixas: i) faixa 1, com as melhores 20 instituições; ii) faixa 2, com as instituições de ensino colocadas entre as posições 21 e 40; iii) faixa 3, com as instituições de ensino colocadas entre as posições 41 e 60; iv) faixa 4, com as instituições de ensino colocadas entre as posições 61 e 79.

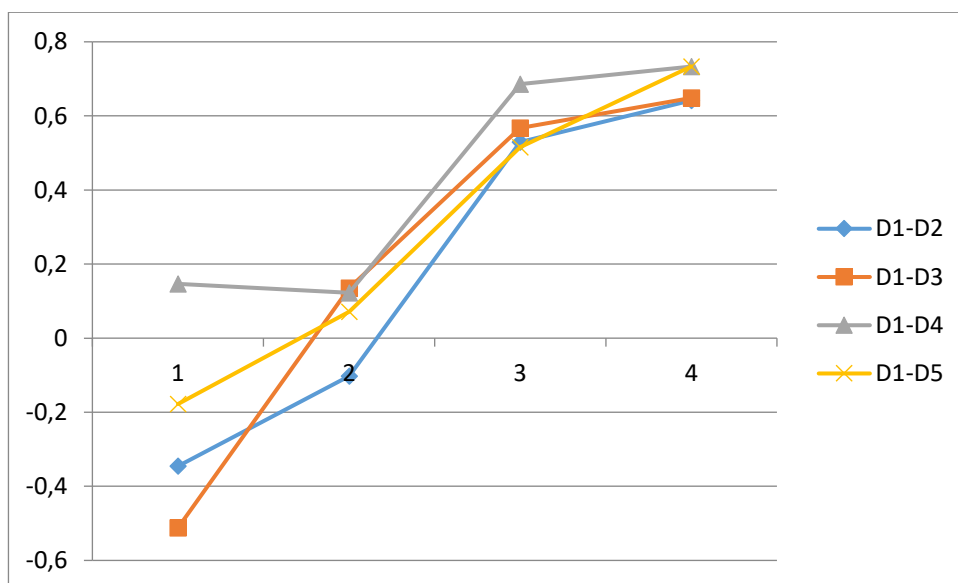
Após isso, é feito um estudo sobre a distribuição de dados reportados para as instituições espanholas para cada faixa em que dividimos as instituições de ensino, com o objetivo de validar as próximas etapas. A Tabela 3 mostra a distribuição de dados reportados para as instituições espanholas para cada faixa de desempenho. Como esperado, a ausência de dados cresce nas faixas de menor desempenho global. Porém, os números não são muito elevados ou discrepantes quando comparados com a média global. A faixa 4, a mais incompleta, ainda assim tem 71,22% de dados disponíveis, contra 54,46% da média mundial. Nota-se também que as duas primeiras faixas têm números muito próximos, 94,06% e 92,19%, respectivamente. Para a faixa 3 temos apenas 7,19% a menos de dados que a faixa 2. Isso significa que a falta de dados não prejudica as análises de correlações feitas no presente trabalho, pelo menos não a ponto de invalidar as conclusões obtidas nessa pesquisa.

Tabela 3 - Distribuição de dados reportados para as instituições espanholas por faixa de desempenho.

Dimensão	Faixa 1	Faixa 2	Faixa 3	Faixa 4
Dados Reportados/Dados Esperados	602/640	590/640	544/640	433/608
Dados Reportados (%)	94,06%	92,19%	85,00%	71,22%

Os resultados obtidos nessas correlações são mostrados abaixo, em forma de gráficos, onde os números 1 a 4 representam as quatro faixas em que as instituições de ensino espanholas foram divididas. Para as Dimensões foi utilizada a nomenclatura: D1: Ensino e aprendizagem, D2: Pesquisa, D3: Transferência de conhecimento, D4: Internacionalização, D5: Envolvimento regional.

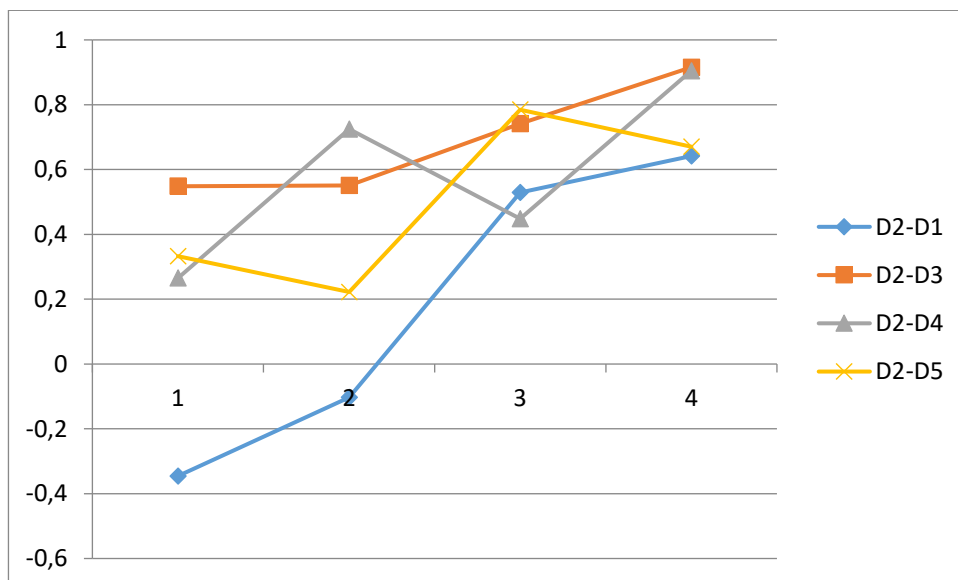
Figura 2 – Correlação da Dimensão D1 (Ensino e aprendizagem) com as demais Dimensões.



A Figura 2 mostra as correlações para a Dimensão D1 (Ensino e aprendizagem). Ela tem correlações negativas na faixa 1 (superior) com as Dimensões D2 (Pesquisa), D3 (Transferência de conhecimento) e D5 (Envolvimento regional), com valor negativo mais elevado (-0,51195) com D3 (Transferência de conhecimento). Esses coeficientes de correlação crescem para as faixas seguintes de forma contínua. Isso significa que, entre as instituições mais bem avaliadas, ter um bom desempenho em Ensino e Aprendizagem implica em uma tendência estatística de ter um pior desempenho nessas outras três Dimensões. Na faixa 2 (instituições colocadas entre as posições 21 a 40) essas correlações ficam próximas de zero, indicando que ter bom desempenho em D1 (Ensino e aprendizagem) e bom desempenho nas Dimensões D2 (Pesquisa), D3 (Transferência de conhecimento) e D5 (Envolvimento regional) são fatos não relacionados. Para a faixa 3 (instituições colocadas entre as posições 41 a 60), a correlação é da ordem de 0,6, se aproximando de 0,8 na faixa 4 (instituições colocadas entre as posições 61 a 79). Isso significa que, na metade inferior de nossa classificação geral, existem fortes correlações positivas e os desempenhos em D1 (Ensino e

aprendizagem) estão muito relacionados aos desempenhos nas Dimensões D2 (Pesquisa), D3 (Transferência de conhecimento) e D5 (Envolvimento regional). Com a Dimensão D4 (Internacionalização), as tendências das correlações são similares, com a diferença de que uma correlação positiva já existe desde a faixa 1 (primeiras 20 colocadas), embora pequena (0,146897). Esse número é bem similar para a faixa 2 (instituições colocadas entre as posições 41 a 60), e depois disso os índices de correlação passam a seguir a mesma tendência de crescimento das demais correlações nas faixas 3 e 4. Então, os desempenhos em D1 (Ensino e aprendizagem) têm uma tendência a gerar desempenhos similares em D4 (Internacionalização), com essa tendência sendo maior para a metade inferior da classificação geral.

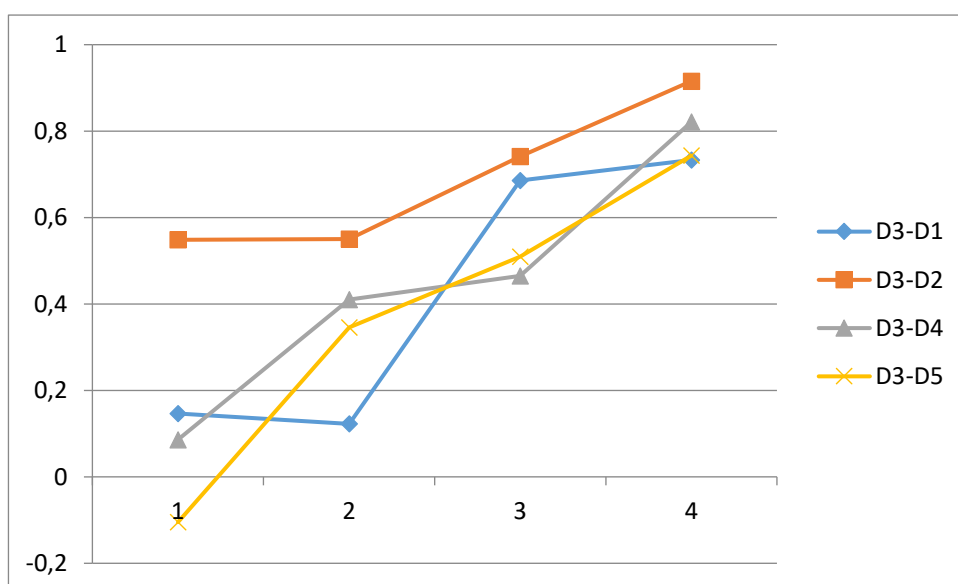
Figura 3 – Correlação da Dimensão D2 (Pesquisa) com as demais Dimensões.



A Figura 3 mostra a evolução dos índices de correlação para a Dimensão D2 (Pesquisa). Essa correlação já foi estudada com relação à Dimensão D1 (Ensino e aprendizagem), mostrando ter uma correlação crescente iniciando com valor negativo na faixa 1 (superior) chegando até um valor um pouco superior a 0,6 na faixa 4. Em relação à Dimensão D3 (Transferência de conhecimento), existe sempre uma correlação positiva, sendo levemente inferior a 0,6 na metade superior da classificação (faixas 1 e 2) e aumentando esse valor para as faixas 3 (quase 0,8) e 4 (acima de 0,9). Isso significa que as instituições tendem a ter desempenhos similares em D2 (Pesquisa) e D3 (Transferência de conhecimento) em todas as faixas, e essa tendência aumenta à medida que examinamos as faixas mais baixas do

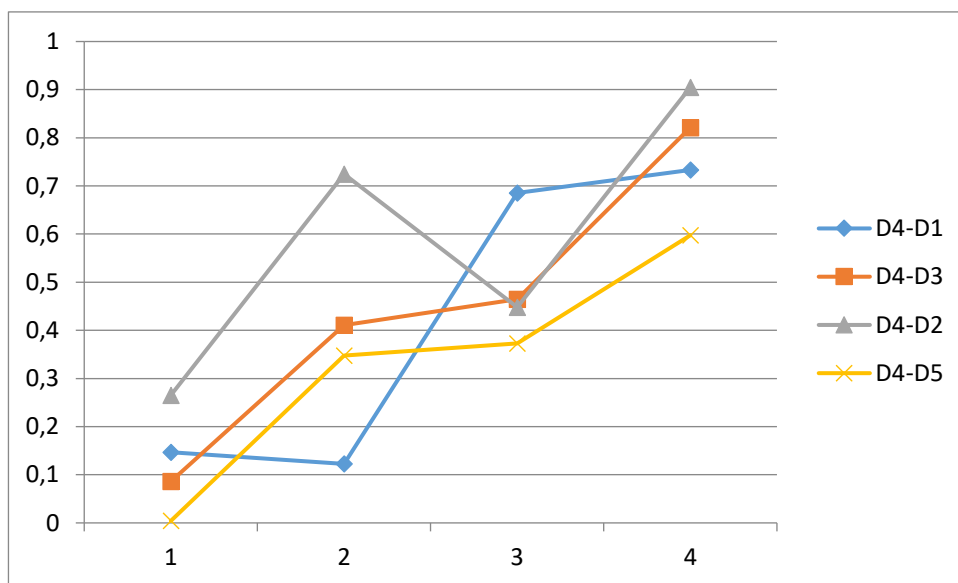
ranking geral. É a correlação mais forte encontrada entre todas as Dimensões das instituições de Ensino espanholas. Em relação à Dimensão D4 (Internacionalização), ela possui sempre uma correlação positiva, porém oscilatória, com máximos locais nas faixas 2 e 4 e mínimos locais nas faixas 1 e 3. Isso implica que instituições com bom desempenho em pesquisa, em geral, são também internacionalizadas. Em relação à Dimensão D5 (Envolvimento regional), existe também uma correlação positiva e oscilatória, mas agora com mínimos locais nas faixas 2 e 4 e máximos locais nas faixas 1 e 3, o que significa que instituições com bom desempenho em pesquisa, em geral, tem também um bom desempenho em Envolvimento Regional.

Figura 4 – Correlação da Dimensão D3 (Transferência de conhecimento) com as demais Dimensões.



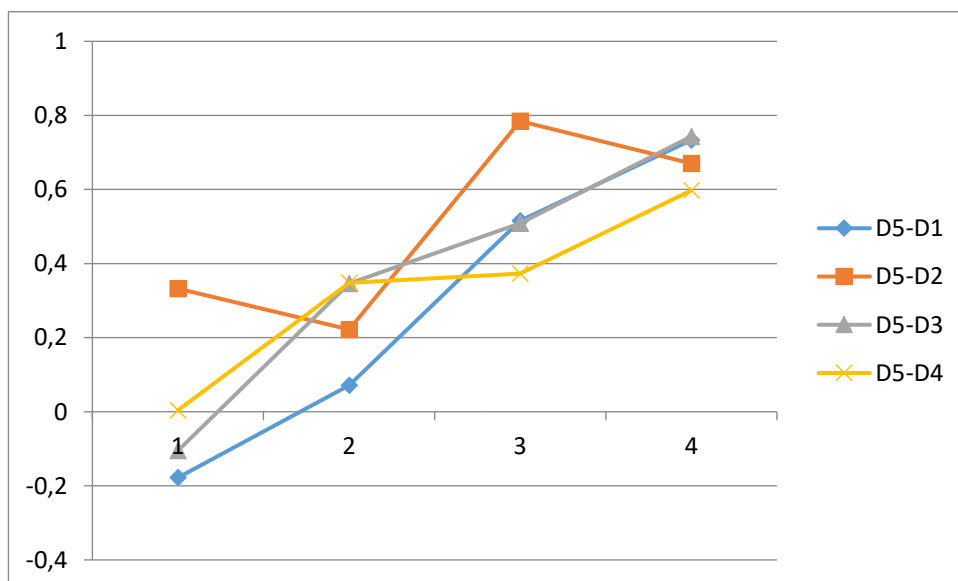
A Figura 4 mostra os resultados para a Dimensão D3 (Transferência de conhecimento). Ela possui, de forma geral, correlações crescentes com todas as demais Dimensões, à medida que avançamos nas quatro faixas estudadas. Com exceção da correlação com D5 (Envolvimento regional), todas as demais correlações são positivas, indicando uma tendência estatística de que os desempenhos em D3 (Transferência de conhecimento) sejam similares nas demais Dimensões, em particular para as faixas mais baixas da classificação geral. A única exceção a essa regra é a correlação entre D3 (Transferência de conhecimento) e D5 (Envolvimento regional), na faixa 1, com um pequeno valor negativo de cerca de -0,1.

Figura 5 – Correlação da Dimensão D4 (Internacionalização) com as demais Dimensões.



A figura 5 faz uma análise para a Dimensão D4 (Internacionalização). Ela também possui, de forma geral, correlações crescentes com todas as demais Dimensões à medida que avançamos nas quatro faixas estudadas. Todas as correlações são positivas, indicando uma tendência estatística de que instituições com melhor desempenho em D4 (Internacionalização) tenham também melhores desempenhos nas demais Dimensões, em particular para as faixas mais baixas da classificação geral. Destaca-se como comportamento diferente a correlação oscilatória entre D2 (Pesquisa) e D4 (Internacionalização), com mínimos locais nas faixas 1 e 3 e máximos locais nas faixas 2 e 4.

Figura 6 – Correlação da Dimensão D5 (Envolvimento regional) com as demais Dimensões



A Figura 6 mostra as correlações entre a Dimensão D5 (Envolvimento regional) e as demais Dimensões. De forma geral, a regra de correlações crescentes à medida que avançamos nas quatro faixas estudadas segue válida. A maioria das correlações é positiva, indicando uma tendência estatística de que instituições com melhor desempenho em D5 (Envolvimento regional) tenham também melhores desempenhos nas demais Dimensões, em particular para as faixas mais baixas da classificação geral. Apenas as correlações entre D1 (Ensino e aprendizagem) e D3 (Transferência de conhecimento) possuem valores negativos na faixa 1, mas com valores absolutos pequenos. Isso mostra uma leve inversão entre os desempenhos em D5 (Envolvimento regional) e D1 (Ensino e aprendizagem) e D3 (Transferência de conhecimento). Destaca-se também, como comportamento diferente, a correlação oscilatória com D2 (Pesquisa), com mínimos locais nas faixas 2 e 4 e máximos locais nas faixas 1 e 3.

Os dados brutos das correlações estão colocados nas Tabelas 4 a 7, mostradas abaixo, para detalhes numéricos mais precisos.

Tabela 4 - Correlação entre as Dimensões do U-multirank para as instituições de Ensino da Espanha na faixa 1 (1-20 colocadas)

	D2	D3	D4	D5
D1	-0,34557	-0,51195	0,146897	-0,17783
D2		0,548515	0,265085	0,332664
D3			0,086038	-0,1049
D4				0,004151

Tabela 5 - Correlação entre as Dimensões do U-multirank para as instituições de Ensino da Espanha na faixa 2 (21-40 colocadas)

	D2	D3	D4	D5
D1	-0,10255	0,135541	0,122548	0,071258
D2		0,550526	0,724639	0,22223
D3			0,410522	0,345923
D4				0,347644

Tabela 6 - Correlação entre as Dimensões do U-multirank para as instituições de Ensino da Espanha na faixa 3 (41-60 colocadas)

	D2	D3	D4	D5
D1	0,529352	0,567769	0,685722	0,515697
D2		0,74143	0,447992	0,784635
D3			0,464888	0,509164
D4				0,373092

Tabela 7 - Correlação entre as Dimensões do U-multirank para as instituições de Ensino da Espanha na faixa 4 (61-79 colocadas)

	D2	D3	D4	D5
D1	0,641833	0,648477	0,733237	0,733572
D2		0,915685	0,905027	0,670534
D3			0,821205	0,743358
D4				0,597564

As estatísticas desenvolvidas para a Espanha foram mais gerais, usando todas as 79 instituições listadas no ranking, e mostraram claramente que as instituições de melhor desempenho (primeiro quartil) apresentam correlações negativas entre várias Dimensões, mas as de desempenho inferiores, mesmo as de segundo quartil, já apresentam correlações positivas com valores maiores, próximas de 1, no terceiro e, principalmente, no quarto quartil. Isso indica que as instituições de menor desempenho geral são mais generalistas, tendo mais regularidade nas diferentes Dimensões e uma maior homogeneidade de desempenho entre as diferentes Dimensões.

Destaca-se, em particular, no caso das instituições de ensino espanholas, a forte correlação entre as Dimensões “Pesquisa” e “Transferência de Conhecimento”, sempre positivas, acima de 0,5, e chegando a valores acima de 0,9 no último quartil.

Análise Preliminar das Dimensões e Indicadores do “U-Multirank” com vistas à utilização na avaliação da Pós-Graduação brasileira

É importante lembrar que o “U-Multirank” foi criado para fornecer informações a futuros alunos, e não fazer uma classificação geral ou mesmo dar notas globais a instituições Acadêmicas. Nesse sentido ele não pode ser simplesmente utilizado para avaliação da pós-graduação brasileira na sua forma atual. Sua utilidade está em fornecer Dimensões e Indicadores que possam ser usados no Brasil, após adequação a nossa realidade. Além disso, deve-se também discutir diferentes formas de combinar essas Dimensões e Indicadores em uma nota final, mesmo que por faixas, como é feito no momento pela CAPES. Existem muitas opções possíveis. Uma sugestão pode ser atribuir pesos a cada Dimensão, com esses pesos variando de acordo com a missão da instituição Acadêmica. Alguns pesos poderiam até mesmo ser zero, implicando no descarte da Dimensão. Isso ajudaria no sentido de levar em conta a missão da instituição no momento da avaliação, que seria sua nota final definida apenas por algumas das Dimensões definidas pela CAPES, aquelas que são compatíveis com sua missão.

Analisando agora os Indicadores presentes no “U-Multitank”, do ponto de vista de seu uso na avaliação da pós-graduação brasileira, é possível fazer alguns comentários iniciais.

Certamente os Indicadores relativos à graduação devem ser eliminados, tais como “Taxa de formandos na graduação”, “Graduação no tempo esperado (bacharéis)”, “Programas de bacharelado em língua estrangeira”, “Formandos na Graduação trabalhando na região” e “Estágios de estudantes na região” (embora não seja um Indicador exclusivo da graduação, essa atividade é rara em pós-graduação).

Existem alguns outros Indicadores que são de difícil aplicação, tais como “Rendas obtidas de desenvolvimento profissional contínuo”, que é pouco comum no Brasil, não se aplica a muitas áreas do conhecimento e tem difícil verificação. “Vagas de pós-doutorado” também não é um Indicador muito adequado, por dependerem de bolsas de agências de fomento, as quais tem pouca disponibilidade, principalmente no momento atual. “Publicações de acesso aberto” também é um Indicador complexo. Em algumas áreas, como as humanidades, existe um estímulo ao seu uso, inclusive com periódicos aceitando, ou até mesmo exigindo que trabalhos submetidos sejam previamente disponibilizados como “preprint” em acesso aberto. Porém, áreas ligadas a ciências exatas ainda não aceitam essa política, e revistas com acesso aberto cobram taxas altas de publicação, muitas vezes superiores a US\$ 2.000,00. Esse indicador traria um forte domínio por instituições com mais recursos financeiros, sem que necessariamente tenham melhor qualidade. “Equipe acadêmica internacional” também é uma questão que deve ser analisada com muito cuidado. Esse Indicador estimula as instituições a evitar políticas endógenas, o que é um fator positivo ao fluxo de idéias e inovações. Mas é preciso também levar em conta que instituições privadas possuem muito mais liberdade na contratação e demissão no corpo docente, podendo alterar mais facilmente sua equipe de trabalho para aumentar o número de docentes internacionais. Nas universidades públicas essa flexibilidade não existe e novas contratações dependem de autorizações para realização de concursos, algo que não está sob controle completo de departamentos, às vezes nem mesmo das instituições. Isso poderia gerar uma desvantagem grande para instituições públicas em relação às instituições de ensino privadas, sem que a razão dessa desvantagem seja baseada em qualidade. “Publicações interdisciplinares” também não é um Indicador muito objetivo, pois a definição de interdisciplinaridade é muito vaga para elaborar critérios claros.

O Indicador do “U-Multirank” “Graduação no tempo esperado” (mestrado, certamente estendido a doutorado) já foi usado pela CAPES em ciclos avaliativos anteriores. Esse indicador está excluído da avaliação referente ao período 2017-2020 devido à pandemia gerada pelo COVID-19, que atrasou muitos trabalhos de forma involuntária. Mesmo em condições normais, existem muitas objeções ao seu uso, pois ele: i) motiva o desenvolvimento de pesquisas mais simples e rápidas, penalizando trabalhos experimentais e aplicados que podem ter um maior tempo de desenvolvimento; ii) gera estratégias que nada contribuem para a melhoria de qualidade da Pós-Graduação brasileira, como matricular no programa apenas alunos que já tenham cursado disciplinas e/ou desenvolvido pesquisas anteriormente, reduzindo assim, artificialmente, os tempos de titulação; iii) eliminar ou reduzir alunos que

sejam também profissionais, reduzindo a ligação instituição de ensino-indústria e elitizando a Pós-Graduação a alunos que não tenham necessidade de trabalhar durante os estudos.

Alguns desses Indicadores já estão sendo utilizados na avaliação 2017-2020, pelo menos em algumas das áreas de avaliação, e que provavelmente serão mantidos, por serem considerados eficientes e verificáveis. Eles são: “Taxa de citação” (por exemplo, a área de Engenharias III utiliza o índice h2, que é o número de docentes com índice h-Scopus maior ou igual a h); “Receitas de pesquisa externa” e “Parcerias de pesquisa estratégica”, pois esses Indicadores medem a capacidade do Programa em obter recursos externos e parceiros para pesquisa, um forte indicativo de qualidade e aplicabilidade do Programa; “Publicações mais citadas”, obtido solicitando aos Programas uma lista das publicações mais citadas de cada docente e o número de citações de cada uma delas em uma dada base de dados internacional como Scopus ou Web of Science; “Spin-offs”, que é avaliado solicitando uma lista dos produtos mais relevantes produzidos em cada Programa. Um ponto a ser lembrado com relação ao índice h2 é o fato dele considerar apenas os docentes com maior índice h, com os demais não tendo nenhuma contribuição. Isso colabora para que os programas tenham jovens docentes, sem que eles prejudiquem a avaliação. Por outro lado, apenas uma parte dos docentes é avaliada nesse Indicador. Uma opção a ser estudada seria considerar a média dos índices h de todos os docentes, ou de metade dos docentes, para permitir a permanência no Programa de docentes com viés tecnológico ou em início de carreira, sem prejudicar a nota do Programa.

Indicadores que podem ser utilizados sem muita dificuldade e que são passíveis de verificação são: “Publicações Profissionais”; “Publicações citadas em patentes”; “Patentes em co-autoria com a indústria”; “Rendas de fontes privadas” e “Empresas criadas por formandos”, que medem a capacidade do Programa em desenvolver novas tecnologias de interesse da indústria ou de gerar novas empresas; “Programas de mestrado em língua estrangeira”, certamente estendido ao Doutorado; “Mobilidade estudantil”; “Publicações conjuntas internacionais”; “Porcentagem de Doutores Estrangeiros formados”, itens que medem a internacionalização do Programa; “Publicações conjuntas regionais”; “Publicações regionais com parceiros industriais”; “Rendas de fontes regionais”; “Formados no mestrado trabalhando na região”, que medem a inserção regional do Programa, certamente estendido ao Doutorado. Uma preocupação com os Indicadores ligados a inserção regional deve ser a de evitar penalizar Programas localizados em regiões de menor índice de desenvolvimento humano. Regiões assim têm poucas indústrias e outras atividades econômicas, e isso deve gerar menores números em Indicadores como “Publicações conjuntas regionais”,

“Publicações regionais com parceiros industriais”, “Rendas de fontes regionais”, “Formados no mestrado trabalhando na região”. Esses Indicadores são importantes, mas é necessário relativizá-los de alguma forma com as atividades econômicas da região.

Indicadores já usados por muito tempo, e que provavelmente continuarão a ser usados: “Taxa de formandos no mestrado”, certamente estendido ao Doutorado; “Publicações (tamanho normalizado)”, que deve ter preferência em relação a “Publicações (números absolutos)”, por considerar o tamanho do Programa de Pós-Graduação; “Taxa de graduação no mestrado” (certamente estendido ao Doutorado), pois mede a eficiência do Programa em formar mestres e doutores, a função primária da Pós-Graduação; “Patentes concedidas” (normalizadas pelo tamanho), que deve ter prioridade em relação a “Patentes concedidas” (números absolutos), por considerar o tamanho do Programa; “Produtos relacionados à arte”, onde aplicável.

Conclusões

O presente trabalho fez uma análise do ranking acadêmico internacional “U-Multitank”, que ajuda a entender como fazer uma avaliação acadêmica multidimensional, em particular enfatizando suas vantagens e desvantagens. O “U-Multirank” cobre um conjunto maior de Indicadores quando comparado com outros rankings acadêmicos, tendo uma visão mais geral das instituições avaliadas. Por outro lado, é preciso um maior esforço na coleta de dados, pois ele tem um alto índice de falta de dados. Outra vantagem desse ranking é que ele usa apenas dados objetivos, sem levar em conta “reputação acadêmica”.

Foi também mostrado que as instituições Acadêmicas espanholas mais bem colocadas no ranking possuem uma correlação mais fraca entre as Dimensões do “U-Multirank”, comparadas com as de desempenho inferior, o que significa que elas têm um comportamento mais heterogêneo entre os pontos medidos.

Outro fato observado foi de que o “U-Multirank”, apesar de ter sido criado para fornecer informações a futuros alunos, sem dar notas globais às instituições avaliadas, pode ajudar na avaliação da Pós-Graduação brasileira, no sentido de fornecer Dimensões e Indicadores que possam ser usados nesse processo. Seria uma forma de aplicar o conceito de avaliação multidimensional utilizado pelo “U-Multirank”.

De forma geral, espera-se que essa nova metodologia possa trazer Programas fortes em qualidade de ensino, inserção regional e ligação com o setor industrial para patamares mais elevados em termos de notas da CAPES. Por outro lado, os programas fortes em pesquisa e internacionalização, como os das áreas relacionadas à ciência, devem competir entre

si, e esse fato deve elevar o patamar numérico requerido para esses Indicadores para obter uma data nota, reduzindo assim as notas de alguns programas, mesmo que mantenham o desempenho de outros períodos. Ou seja, deverão ocorrer mudanças significativas nas notas de muitos Programas, em particular os pertencentes ao campo da ciência e que estejam na faixa de notas 4 e 5, as intermediárias da tabela de notas.

Também se pode observar que esse método multidimensional pode ser estendido para a graduação, como é feito na Europa, e poderíamos ter métodos similares para avaliar a graduação e a pós-graduação brasileiras, o que poderia ser bastante interessante do ponto de vista de uniformização de qualidade entre os dois níveis. Além disso, muitos dados seriam os mesmos, e isso poderia facilitar o trabalho das duas avaliações.

Do ponto de vista de correlação entre notas obtidas no sistema atual e no novo sistema proposto, sabemos que a avaliação atual da CAPES é bastante focada em pesquisa e internacionalização, então se espera uma correlação maior entre as notas atuais e as notas obtidas nessas duas Dimensões.

De forma geral, todos os programas devem se preocupar e se preparar para novos Indicadores e formas de avaliação, mas uma primeira análise indica que cursos das áreas ligadas à ciência podem estar entre os mais afetados pelas mudanças e com maior probabilidade de redução de notas.

Referências

ANDRIOLA, W. B.; ARAÚJO, A. C. Uso de indicadores para diagnóstico situacional de instituições de Ensino Superior. **Ensaio: aval. pol. públ. Educ.**, Rio de Janeiro, v. 26, n. 100, p. 645-663, jul. 2018. Disponível em: <https://doi.org/10.1590/s0104-40362018002601062>. Acesso em: 20 jul. 2021.

BARREYRO, G. B. De exames, rankings e mídia. **Avaliação**, Campinas; Sorocaba, v. 13, n. 3, 2008. Disponível em: <https://doi.org/10.1590/S1414-40772008000300017>. Acesso em: 20 jul. 2021.

CALDERÓN, A. I.; FRANÇA, C. M. Os rankings acadêmicos da educação superior: apontamentos no campo da avaliação educacional. *In*: ROTHEN, J. C.; SANTANA, A. C. M. (org.). **Avaliação da Educação: referências para uma primeira conversa**. São Carlos: Ed. Universidade de São Carlos, 2018. p. 95-114.

CASTRO, C. M.; SOARES, G. A. D. Avaliando as avaliações da CAPES. **Revista de Administração de Empresas**, São Paulo, v. 23, n. 3, p. 63-73, set. 1983.

CAPES -COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. **Edital nº. 41/2017 - Programa Institucional de Internacionalização (CAPES-PrInt)**. Brasília-DF: CAPES, 2017.

CAPES - COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Avaliação Multidimensional de Programas de Pós-Graduação. Relatório DAV**. Brasília-DF: CAPES, 2019.

DRESCH, J. F. Avaliação da educação e o cenário midiaticizado da responsabilização. *In*: ROTHEN, J. C.; SANTANA, A. C. M. (org.). **Avaliação da educação**: referências para uma primeira conversa. São Carlos: Ed. Universidade de São Carlos, 2018. p. 81-94.

FONSECA, M.; OLIVEIRA, J. F. A avaliação da pós-graduação e o seu impacto no trabalho acadêmico dos professores brasileiros. *In*: OLIVEIRA, J. F.; CATANI, A. M.; FERREIRA, N. S. C. (orgs.). **Pós-graduação e avaliação**: impactos e perspectivas no Brasil e no cenário internacional. Campinas: Mercado de Letras, 2010. p. 109-130.

FURTADO, H. L.; HOSTINS, R. C. L. Avaliação da pós-graduação no Brasil. **Revista de Educação PUC-Campinas**, Campinas, v. 19, n. 1, p. 15-23, jan./abr., 2014.

KAWASAKI, B. C. Critérios da avaliação CAPES para Programas de Pós-Graduação. **Revista Adusp**, São Paulo, p. 102-117, maio 2017.

LLOYD, M. W.; ORDORIK, I.; RODRÍGUEZ-GÓMEZ, R. Los rankings internacionales de universidades: su impacto, metodología y evolución. **Cuadernos de Trabajo de la Dirección General de Evaluación Institucional**. México: UNAM, 2011. n. 7.

MENDES, A. L. F.; DUTRA, N. L. L. Índice geral de cursos e qualidade na educação superior: o caso das IES do Estado da Bahia. **Avaliação**, Campinas; Sorocaba, v. 24, n. 1, 2019. Disponível em: <https://www.scielo.br/j/aval/a/bMnYSvkTCKyJcht39bNDB9b/?lang=pt>. Acesso em: 20 jul. 2021.

MOURA, B. A.; MOURA, L. B. A. Ranqueamento de universidades: reflexões acerca da construção de reconhecimento institucional. **Acta Scientiarum. Education**, Maringá, v. 35, n. 2, p. 213-222, 2013.

NOBRE, L.N.; FREITAS, R.R. A evolução da pós-graduação no Brasil: histórico, políticas e avaliação. **Brazilian Journal of Production Engineering (BJPE)**, São Mateus, v. 3, n. 2, p. 18-30, 2017.

PATRUS, R.; SHIGAKI, H. B.; DANTAS, D. C. Quem não conhece seu passado está condenado a repeti-lo: distorções da avaliação da pós-graduação no Brasil à luz da história da CAPES. **Cadernos EBAPE.BR**, Rio de Janeiro, v. 16, n. 4, p. 642-655, dez. 2018.

PILATTI, L. A.; CECHIN, M. R. Perfil das universidades brasileiras de e com potencial de classe mundial. **Avaliação**, Campinas; Sorocaba, v. 23, n. 1, 2018. Disponível em: <https://doi.org/10.1590/S1414-40772018000100006>. Acesso em: 20 jul. 2021.

PRADO, A.F.B.A.; AGUILAR, L. E. Three models of international academic ranking: comparative analysis. *In*: CONGRESO IBEROAMERICANO DE EDUCACIÓN COMPARADA, 3., Montevideo, Uruguai, 2021.

PRADO, A.F.B.A. Performances of the Brazilian Universities in the “U-MULTIRANK” in the Period 2017-2020. **SciELO Preprint**, São Paulo, 2021a. Preprint. Disponível em: <https://preprints.scielo.org/index.php/scielo/preprint/view/2351>. Acesso em: 20 jul. 2021.

PRADO, A.F.B.A. Multidimensional Academic Evaluation using the 'U-Multirank'. **SciELO Preprint**, São Paulo, 2021b. Preprint. Disponível em: <https://preprints.scielo.org/index.php/scielo/preprint/view/2661/version/2813>. Acesso em: 20 jul. 2021.

- PRADO, A.F.B.A. Study of the multidimensional academic ranking “U-Multirank” **Scielo Preprint**, São Paulo, 2021c. Preprint. Disponível em: <https://preprints.scielo.org/index.php/scielo/preprint/view/3063>. Acesso em: 20 jul. 2021.
- RIGHETTI, S. Quando a sociedade é quem avalia: a experiência do Ranking Universitário Folha e os indicadores de ensino superior. **COMCIÊNCIA (UNICAMP)**, Campinas, v. 166, p. 1, 2015.
- ROTHEN, J. C.; SANTANA, A. C. M. (org.). **Avaliação da educação: referências para uma primeira conversa**. São Carlos: Ed. Universidade de São Carlos, 2018.
- SHIN, J. C.L.; TOUTKOUSHIAN, R. K.; TEICHER, U. **University rankings, theoretical basis, methodology and impacts on global higher education**. New York: Springer, 2011.
- SOUZA, V. C. Qualidade na educação superior: uma visão operacional do conceito. **Avaliação**, Campinas; Sorocaba, v. 22, n. 2, 2017. Disponível em: <https://www.scielo.br/j/aval/a/3d5tt9h3MSx4Sfq4s35WSCN/?lang=pt>. Acesso em: 20 jul. 2021.
- STACK, M. **Global University Rankings and the Mediatization of Higher Education**. Palgrave Studies in Global Higher Education. England: Macmillan Publishers Limited, 2016.
- THERY, H. Classificações de universidades mundiais, “Xangai” e outras. **Estudos Avançados**, São Paulo, v. 24, n. 70, p. 185-205, 2010.
- VAN VUGHT, F. A.; ZIEGELE, F. (ed.). **Multidimensional ranking: the design and development of U-Multirank**. New York: Springer, 2012.