ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 09, Issue, 01, pp.25526-25532, January, 2019

ORIGINAL RESEARCH ARTICLE

OPEN ACCESS

AN ANALYSIS OF SIGNIFICANT FACTORS IN GLOBAL SCHOOL RANKINGS

Wanmin Li, Yucao Che and *Paul D. Berger

Bentley University, Waltham, MA 02452, U.S.A.

ARTICLE INFO

Article History:

Received 26th October, 2018
Received in revised form
10th November, 2018
Accepted 22nd December, 2018
Published online 30th January, 2019

Key Words:

University Rankings, ANOVA,
Interaction, Orthogonal Breakdown of SSQs,
Latin-square Designs.

ABSTRACT

With the rapid development of the education industry and the trend of globalization, World University Ranking has become an important index for educational institutions and students. When students choose which schools they wish to attend, they often refer to the ranking list. Moreover, recruiters have preference for candidates from target schools, which are defined by rankings to some extent. School/university ranking also becomes a catalyst for policy change or administration implementation in a university. In this paper, we perform statistical analyses to determine the significant factors that affect university rankings.

Citation: Wanmin Li, Yucao Che and Paul D. Berger, 2019. "An Analysis of Significant Factors in Global School Rankings", *International Journal of Development Research*, 09, (01), 25526-25532.

INTRODUCTION

Ranking is an important factor for most schools of higher learning/universities. Although some schools claim they do not put strong consideration on its ranking, more attention has been paid to rankings nowadays. On the one hand, the rankings have been the reason for deans getting fired, and for a school's changing policies (How Important is a School's Ranking, 2010); on the other hand, employers tend to recruit candidates from "target schools," the universities that rank top in business, technology, or other specific fields. However, ranking universities is a difficult and complicated practice which sometimes involve controversies and political issues. There are many different national and international university-ranking systems and many of them are different, or even disagree with one another. The data used in this paper were collected from Kaggle.com, and the dataset contains global university-rankings from *Times Higher Education World University Rankings.* The data show the world's top 800 universities ranking in 2016 and were designed to identify the best universities in the world under the fairest evaluation process. There are a total of 800 schools in the dataset, but only the 200 schools with the top rankings have complete total scores on all of the input factors recorded.

Corresponding author: **Paul D. Berger,**
Bentley University, Waltham, MA 02452, U.S.A.

In our paper, we first calculate the required sample size (so that the results have a certain *power*,) and then perform a one-factor ANOVA to determine if certain factors have a significant influence on a university's ranking. We do this for three factors, Research, Teaching, and Location. We then conduct a two-factor ANOVA to test if there is an interaction effect between Research and Teaching. We then perform a one-factor ANOVA to study the impact of Location (variables defined in the next section.) Finally, we study the three factors, Number of Students, Student-Staff ratio and Number of citations of research using a Latin-Square Design. For some of the factors that are shown to be significant, we do a more detailed analysis by conducting an orthogonal breakdown of sums-of-squares, to determine specific differences among the levels of the factor that are contributing to the variability of university ranking. The analysis has potential to be particularly meaningful for the universities who want to improve their rankings and improve their overall education quality.

LITERATURE REVIEW

The world university ranking agency ranks universities based on the values of predetermined factors that are thought to reflect the academic excellence of academic institutions. Without doubt, Research and Teaching would be expected to be included.

However, there are also a variety of other factors that are used for ranking; for example, Location is an aspect of internationalization of higher education spurred by globalization (Tan, Goh., 2014). Their studies indicate the fact that Western Europe and North America are the world regions of choice for international students, and it is a daunting task for educational institutions from developing countries to attract these students. The Student-Staff ratio factor (SSR) is also included. SSRs are seen as a proxy for quality, and educational institutions tend to work intuitively toward lower SSRs (McDonald, 2013). The World University rankings have become a hot topic in the education industry, and there is not full agreement about the validity of the factors that have been adopted for measurement, and the accuracy of the scoring system. The inescapable feature of the world university-ranking system is to use to a weigh-and-sum to process the data (Soh, 2017). Soh pointed out that the weigh-and-sum approach assumes that the indicators are mutually compensatory: "*a low score for, say, citations, can be compensated by a high score for, say, proportion of international students.*" However, the "weight discrepancies" issue has been found to not follow the assigned weighting schemes (Soh, 2017). "Indicator redundancy" happens because some factors are highly correlated (Soh, 2017). Moosa (2018) mentioned the concept of the "Cinderella Effect," that research is often accorded a higher status, because publications result in a distinct increase on ranking, and thus teaching, which is less easy to measure, takes on lesser importance.

The studies of Schwekendiek (2015) demonstrate the similar phenomena that the universities in Germany with good reputations lag behind in the world ranking because German schools tend to hire only a few permanent scholars to keep the taxes low for free university education, and this fact results in fewer publications. German scholars, whatever their age may be, do not necessarily publish in internationally acknowledged journals such as those indexed in the SSCI, SCI, A&HCI, or SCOPUS index. Many European countries offer higher education with free tuition and face the same problems. The controversy about whether student-staff ratio (SSR) has a significant effect on student achievements has continued; Glass and Smith (1979) reported that class size affects the quality of the classroom environment, while Shane (1961) concluded that, using academic achievement as a criterion, there were no differences between large and small classes across 32 studies. This paper tests the significance of selected factors, to address some questions mentioned, such as whether the factors that are claimed to have a heavy weight, such as "teaching," and "SSRs" are truly significant factors for the university rankings. Also, we want to test whether the scores for ranking are different for various locations; for example, North America and Europe attracted a great number of international students, and, yet, universities in countries such as Germany, with different education policies, tend to have slightly lower rankings.

MATERIALS AND MEHTODS

The dataset that is used for analysis is the World University Rankings (Times Higher Education) in 2016. The dataset, downloaded from Kaggle, has a total of 800 observations available. However, there are empty or incomplete data on key features for universities ranked above the Top 200; therefore only the top 200-ranked universities are considered. The dependent variable for our experiment is the "Total Score" of the university. The six factors considered are the "levels" of *Research, Teaching, Location, Number of Students, Student-Staff ratio,* and *Number of Citations.* These six factors will be tested in various ways to determine whether they are significant factors for university ranking. For the experiment designs, each factor will have several "levels." The setting of the levels is based on the data distribution. We have the following variables/factors:

- ***"Total Score"*** – Total score for university, used to determine rank – (Y) - our dependent variable.
- ***"Research"*** – Measurement of research volume.
- ***"Teaching"*** – Measurement of learning environment; how committed an institution is on nurturing the next generation of academics.
- ***"Location"*** – The University's location - which country / area.
- ***"Citation"*** – Measurement of research's influence; "how much each university is contributing to the sum of human knowledge: they tell us whose research has stood out, has been picked up and built on by other scholars and, most importantly, has been shared around the global scholarly community to expand the boundaries of our understanding, irrespective of discipline"
- ***"Number of students"*** – The total number of students accommodated by the school.
- ***"Student-Staff Ratio"*** (SSR) - The number of students who attend a school or university divided by the number of teachers in the institution; often used as a proxy for class size.

We use ANOVA tests and perform one-factor, two-factor and three-factor designs. The significance level α is always .05, and the power of test, $1 - \beta$, is desired to be 0.95. The conclusions will be based on the p-value from the ANOVA test. Additionally, we also perform, in several instances, the Kruskal-Wallis (KW) test (as a backup, in case some parts of the data set were "very non-normally distributed.") If the p-value < .05, the null hypothesis will be rejected, if the p-value > = .05, null hypothesis will be accepted. The null hypothesis for the ANOVA (and, essentially, the KW) tests is there is no difference on university rank due to the level of the studied factor. In each case, before the ANOVA testing, the desired (really: minimum) sample size was calculated based on the range of the levels ($\triangle$ = Max μ – Min μ), the significance value of .05, power of test ($1-\beta$) of .95, the number of levels, and the ratio ($\triangle/\sigma$). For each factor, the number of replications at each level was equal to or higher than the desired (required) sample size. As we shall describe, due to the limited number of observations in the dataset, we needed to lower the power of test in some cases.

RESULTS AND DISCUSSION

Research Factor: The Research Volume factor is separated into 4 levels. The probability distribution of research scores and the corresponding histogram is shown in Figure 1. The p-value is = .000 from the one-factor ANOVA results shown in Table 1, and for the KW test results shown in Table 2, the p-value also = .000 (rounded to three digits, as all p-values reported in this paper.) This suggests that the Research Factor is a highly significant factor to indicate the school ranking.

Research	Avg	N
< = 40	52.6210526	57
40-60	57.5243243	74
60-80	70.002381	42
80-100	85.4703704	27
		200
Range △	33	
sd	12.0405554	
△/sd	2.72822279	

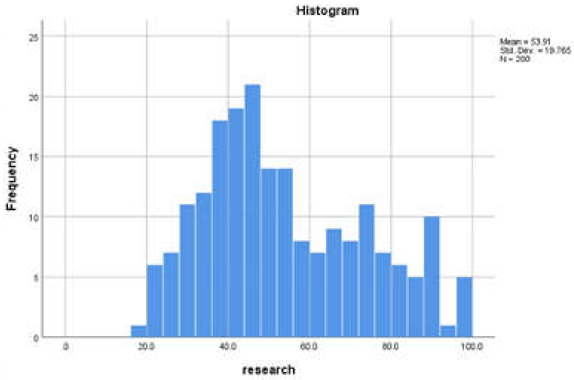

Figure 1. Research Factor Frequency Distribution and Histogram

Table 1. ANOVA Test of Research Factor

Tests of Between-Subjects Effects					
Dependent Variable: TotalScore					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	7502.974[a]	3	2500.991	137.906	.000
Intercept	253214.081	1	253214.081	13962.380	.000
Research	7502.974	3	2500.991	137.906	**.000**
Error	1015.585	56	18.135		
Total	261732.640	60			
Corrected Total	8518.559	59			

Table 2. KW Test Results of Research Factor

Test Statistics[a,b]	
	V1
Chi-Square	50.821
df	3
Asymp. Sig.	**.000**

a. Kruskal Wallis Test
b. Grouping Variable: V2

Teaching	Avg	N
< = 40	53.1194	67
40-60	59.98554	83
60-80	75.33143	35
80-100	88.64	15
		200
Range △	36	
sd	12.04056	
△/sd	2.95008	

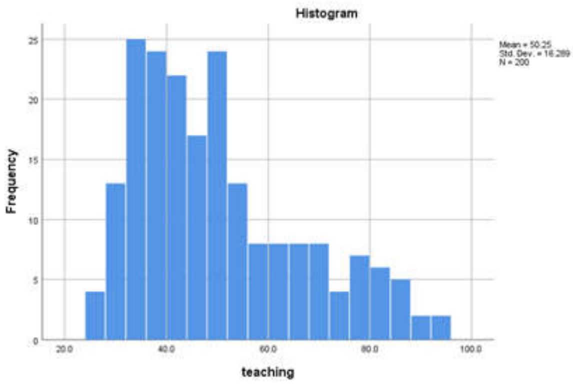

Table 3. ANOVA Test of Teaching Factor

Tests of Between-Subjects Effects					
Dependent Variable: TotalScore					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	10925.758[a]	3	3641.919	105.635	.000
Intercept	291778.214	1	291778.214	8463.128	.000
Teaching	10925.758	3	3641.919	105.635	**.000**
Error	1930.679	56	34.476		
Total	304634.650	60			
Corrected Total	12856.436	59			

Table 4. KW Test Results of Teaching Factor

Test Statistics[a,b]	
	V1
Chi-Square	48.642
df	3
Asymp. Sig.	**.000**

a. Kruskal Wallis Test
a.Grouping Variable: V2

Table 5. Teaching and Research 2-Factor ANOVA SPSS Output

Tests of Between-Subjects Effects					
Dependent Variable: TotalScore					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	20476.415[a]	3	6825.472	159.763	.000
Intercept	264409.203	1	264409.203	6188.996	.000
Teaching	1166.539	1	1166.539	27.305	**.000**
Research	3062.199	1	3062.199	71.677	**.000**
Teaching * Research	162.081	1	162.081	3.794	**.053**
Error	8373.605	196	42.722		
Total	810600.100	200			
Corrected Total	28850.020	199			

Table 6. ANOVA Test of Location Factor

Tests of Between-Subjects Effects					
Dependent Variable: TotalScore					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1460.433[a]	3	486.811	3.182	.026
Intercept	306134.243	1	306134.243	2001.132	.000
Location	1460.433	3	486.811	3.182	**.026**
Error	22794.100	149	152.981		
Total	648576.770	153			
Corrected Total	24254.533	152			

Kruskal Wallis Test

Table 7. KW Results for Location Factor

Test Statistics[a,b]	
	TotalScore
Chi-Square	8.856
df	3
Asymp. Sig.	**.031**

a. Kruskal Wallis Test
a.Grouping Variable: Location

Table 8. Mean Teaching Score for Different Levels

Poor	Average	Good	Excellent
0-40	40^{+}-60	60^{+}-80	80^{+}-100
53.84	60.69	75.77	88.64

Table 9. Orthogonal Matrix of Teaching Score for Different Levels

	Poor	Average	Good	Excellent	Z	Z^2*15
	53.84	60.69	75.77	88.64		
Poor vs. Non-poor	$3/\sqrt{12}$	$-1/\sqrt{12}$	$-1/\sqrt{12}$	$-1/\sqrt{12}$	-18.354	5053.02
Average vs. Good	0	$-1/\sqrt{2}$	$1/\sqrt{2}$	0	10.668	1707.06
Average & Good vs. Excellent	0	$-1/\sqrt{6}$	$-1/\sqrt{6}$	$2/\sqrt{6}$	16.665	4165.58

Table 10. ANOVA Table of Teaching Scores Orthogonal Breakdown of Sum-of-Squares

	Source	SSQ	df	MSQ	F
Between Columns	Poor vs. Non-poor	5053.02	1	5053.02	146.55
	Average vs. Good	1707.06	1	1707.06	49.51
	Average & Good vs. Excellent	4165.68	1	4165.68	120.81
Within Columns		1930.68	56	34.48	

Table 11 Questions Chosen and F-statistics for Orthogonal Breakdown of SSQ's for Location

	F_{calc}
European vs. Others	**5.101**
Asia vs. Oceania	0.001
Asia & Oceania vs. North America	2.846

Table 12. Frequencies for the Different Levels of the 3 Factors

Between-Subjects Factors		
		N
StudentNum	1	30
	2	18
	3	18
StudentStaff	1	30
	2	20
	3	16
Citation	1	19
	2	14
	3	33

Table 13. Student Number/Student-Staff Ratio/Citation ANOVA - SPSS Output

Tests of Between-Subjects Effects Dependent Variable: TotalScore					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3039.020[a]	6	506.503	3.479	.005
Intercept	231655.845	1	231655.845	1591.079	.000
StudentNum	193.250	2	96.625	.664	.519
StudentStaff	66.808	2	33.404	.229	.796
Citation	2386.897	2	1193.448	8.197	**.001**
Error	8590.207	59	145.597		
Total	297096.020	66			
Corrected Total	11629.227	65			

Teaching Factor: The Teaching factor was also separated into 4 levels. The probability distribution of research scores and the corresponding histogram is shown in Figure 2. The p-value = .000 from both the one-factor ANOVA results shown in Table 3, and the KW test shown in Table 4. The p-value = .000 suggests that the Teaching factor is (also) a highly significant factor to indicate university ranking. Although some people believe that Teaching factor is less emphasized in university rankings, the result here indicates that the Top tier schools in the university ranking system have a high Teaching score.

Interaction effect between Teaching & Research: We have noted that both Research and Teaching are significant factors for determining university ranking. We next conducted a two-factor ANOVA to study the interaction effect between Research and Teaching; we are aware that this 2-factor ANOVA somewhat renders redundant the one-factor ANOVA results, but wanted to do perform both sets of tests, given that these two factors are clearly very key. The output is shown in Table 5. The mains effects of research and teaching are significant (each p-value = .000.) The interaction effect has a p-value = 0.053, slightly higher than $\alpha = 0.05$. Therefore, *technically*, we fail to reject the null hypothesis that there is no interaction effect; however, the value being so close to .05 suggests that we consider strongly the possibility that there is interaction between the two factors. If we examine the means in a 2x2 table of "higher" (H) and "lower" (L) Research score, vs. "higher" (H) and "lower" (L) Teaching score (with cutoff point 60), we find that the direction of the interaction is that the impact on university ranking of having a higher Research score is even *higher* when the university's Teaching score is higher than when the university's Teaching score is lower (or conversely, but equivalently, the impact on university ranking of a higher Teaching score is even *higher* when the university's Research score is higher than when the university's Research score is lower.)

Location Factor: The location factor (i.e., location of the university) will be separated into 4 levels – North America, Europe, Oceania (Australia and New Zealand) and Asia. The required sample size for the location factor in a four-level test with significance level of .05 and power of .95 is large. The number of observations available in Oceania and Asia failed to satisfy the required sample size, due to the fact that there are a relatively small number of schools from these two areas within the top 200 of the world universities. Thus, the power of this test is .70. The ANOVA results are presented in Table 6. The p-value = 0.026 from the one-factor ANOVA results is lower than the significance level of .05, indicating that we conclude a significant difference in university ranking for different locations. The Kruskal Wallis test results are shown in Table 7, and the p-value = 0.031, which is slightly higher than the F-test p-value, but not materially so. The universities in North America generally have higher total scores (mean = 67.2) compared to the rest of area in the world (mean = 61.1), and average total score of Europe (60.5) is slightly (but not statistically significantly) lower than Asia & Oceania (mean = 62.6). As earlier noted, different countries structure their education resources differently than other countries (e.g., Germany), and the significance of the location factor seems to further confirm this point, since a lot of European countries conduct their resource allocation for educational institutions in a similar way to that of Germany.

More detailed inquiries using orthogonal breakdowns of sum-of-squares

Teaching: The ANOVA and KW tests lead to the conclusion that university ranking differed by the level (score) of Teaching; however, it did not inform us about the differences *in detail* – only that all four levels of teaching score do not suggest the same university ranking. One way, among many, to inquire about the differences in a more "micro" way, is to break down the sum-of-squares associated with teaching score into orthogonal components, each addressing a more detailed inquiry into the data's message. We sorted the Teaching scores of the first 200 ranked universities. We first displayed a descriptive analysis summary of the Teaching scores data. The distribution and mode can be seen more clearly from the histogram (Figure 2). From Figure 2, we saw that the distribution of the Teaching scores was somewhat skewed to the right, and most of the universities have Teaching scores from 32 to 52. Based on the descriptive statistics, we categorized the top 200 ranked universities into 4 groups/levels and labeled them as indicated –

- <= 40 (poor)
- 40^+ - 60 (average)
- 60^+ - 80 (good)
- 80^+ - 100 (excellent)

With four levels of the factor, we can ask 3 [orthogonal] "micro" questions in order to analyze how Teaching scores specifically influence university ranking, such that the sums-of-squares of the micro questions add [exactly] to the sum-of-squares of the overall sum-of-squares due to Teaching, and the 3 questions are statistically independent. For a more complete discussion of an orthogonal breakdown of sums-of-squares, see, for example, Berger et al., 2018. The 3 micro questions we chose are –

- How different is the average ranking between universities with a poor Teaching score and universities with a non-poor Teaching score (average score or higher)?
- How different is the average ranking between a university an average Teaching score and a university with a good Teaching score?
- How different is the average ranking between a university with an average or good Teaching score and a university with an excellent Teaching score?

To answer these questions, we first calculate the column mean of each level, shown in Table 8. In Table 9, we show the orthogonal (actually, "orthonormal") matrix for the 3 questions chosen. Finally, we display the ANOVA results in Table 10 and consider that $F(1, 56) = 4.00$. From Table 10, we can conclude that there is a significant difference in university ranking between poor and non-poor mean Teaching scores, between average and good mean Teaching scores, and between the average & good vs. excellent mean Teaching scores. In each case the *direction* of the difference is the obvious one – "Non-poor" suggests higher average university ranking than "Poor," "Good" suggests higher average university ranking than "Average," and "Excellent" suggests higher average university ranking than "Good" & "Average."

Research: Next, we performed the same analysis on Research score, and investigated in more detail that how the different level of research quality affects the average university ranking. As with the above Teaching score analysis, we divided the schools into the same four groups, labeled *poor, average, good, excellent* by their Research scores. Then we found the column means for each Research level. We proposed the same three questions as for the Teaching-score analysis. In essence, we got the exactly same results as we obtained for the teaching scores.

Location: In the location one-factor design, we get the p-value = 0.026, meaning that there is significant difference in university ranking among the four different locations (see Table 6.) We earlier indicated that certain countries' rankings in Europe are, in a sense, underestimated. Germany was specifically mentioned before. However, low-cost and free education is available in many European countries. Countries that are providing free or low-cost education also include France, Norway, Iceland, Denmark, Sweden, Finland, Austria, Belgium, Czech Republic, Greece, Italy, and Spain. In this context, our orthogonality analysis investigated the following three questions, noted in Table 11, along with the resulting F-statistics for the orthogonal breakdown of sums-of-squares (SSQ's). The F- table value at $\alpha = .05$ is 4.15. Thus, we note that there is a significant difference between European and non-European countries, but not between Asia and Oceania, nor between Asia & Oceania and North America.

Number of students, Student-staff ratio, and Citations: With the assumptions of no interaction effects existing among the three selected factors, we now use a Latin-Square design to test whether the Number of Students, the Student-Staff ratio and Number of Citations are significant factors for university ranking. The frequencies for the levels of each factor are shown in Table 12, while the ANOVA test results for the Latin-Square design presented in Table 13 suggest that the Number of Students and Student-Staff ratio are not significant factors while the Number of Citations is a significant factor for university rank. That the number of citations is significant is not a surprise, for it is a factor often thought of as a surrogate for the *impact* for the research output on other researchers and the field of endeavor.

Conclusion

Without repeating all of the details of our findings, we simply note that out of our six factors we studied, four of them were found to be significantly related to university ranking in the "common sense" direction. These are Research, Teaching, Location, and Number of Citations. We also found a significant interaction effect between Research score and Teaching score – when either is higher, there is a more positive impact on university ranking when the other one is higher. We also found that the Number of Students and Student-Staff ratio did not show a significant relationship to university ranking. We also provided, in some instances, a detailed breakdown of the differences among the levels of the factors, using an orthogonal breakdown of sums-of- squares.

Limitations and directions for future research: Of course, there are some limitations of this study. First of all, due to missing values of some key factor scores, only 200 observations are used in the study. The results may be different if more observations had been able to be used for analysis, especially since some of the p-values are not so different from .05.

In addition, we set 2 or 3 or 4 levels for each factor depending on the sample distribution, and the results could possibly change if we had chosen a different number of levels and different cutoff-points for the levels. For the three factors study, we used a Latin-Square design, with the assumption that there are no interaction effects among the three factors, but it is possible that interactions are present. The existence of a non-zero interaction effect (which we assumed away) would not invalidate the significance of *Citations*, but could theoretically lead to significance of *Number of Students* and *Student-Staff ratio*, by currently masking the fact that there should be a much lower error term. We do not believe that this is the case (the respective p-values are not close to .05, indeed, being above .50,) but we cannot rule this possibility out with complete certainty. When we performed analysis for location, the number of available data points for each level of location failed to meet the required sample size for 95% power (given that we decided that we would have the same number of schools at each level), so we had only 70% power. However, the location factor was significant, so that, in this case, the point is, essentially, moot. For future research, one could plan to include more observations from different years to increase the power and accuracy of the analyses. And, as we've discovered significant factors that affect school rankings, a next study of interest could be to build a predictive model using these significant factors.

REFERENCES

Berger, P., Maurer, R., and Celli, G. 2018. Experimental Design, with Applications in Management, Engineering, and the Sciences, 2nd Edition, *Springer, Inc*., Switzerland.

Cooper, S. 2014. Research Explains The Importance Of Ranking In The Top 10. Retrieved from https://www.forbes.com/sites/stevecooper/2014/01/29/research-explains the-importance-of-ranking-in-the-top-10/#6c5d299021f4

How important is a school's ranking? (n.d.). Retrieved from http://www.nationaljurist.com/prelaw/how-important-schools-ranking

McDonald, G. 2013. Does size matter? The impact of student–staffratios. *Journal of Higher Education Policy and Management,35*(6),652-667. doi:10.1080/1 360080x. 2013. 844668

Moosa, R. 2018. World University Rankings: Reflections on Teaching and Learning as the Cinderella function in the South African Higher Education System. *African Journal of Business Ethics,12*(1). doi:10.15249/12-1-165

Schwekendiek, D. 2015. Recent changes in World University Rankings: An explorative study of Korea and Germany. *Asia Europe Journal, 13*(4), 361-377. doi:10.1007/s10308-015 0411-4

Shane, H. G. 1961. What research says about class size and human development. *NEA Journal, 50*, 30–32.

Smith, M. L and Glass, G. V. 1979. *Relationships of class-size to classroom processes, teachersatisfaction and pupil affect: A meta-analysis*. San Francisco, CA: Far West Laboratory for Educational Research and Development.

Soh, K. 2016. The seven deadly sins of world university ranking: A summary from several papers. *Journal of Higher Education Policy and Management,39*(1), 104-115 doi:10.1080/1360080x.2016.1254431

Tan, Y. S and Goh, S. K. 2014. International students, academic publications and world university rankings: The impact of globalisation and responses of a Malaysian public university. *Higher Education,68*(4), 489-502. doi:10.1007/s10734-014-9724-2

Tucker, L. 2018. Where Can You Study Abroad for Free? Retrieved from https://www.topuniversities.com/ student-info/studying-abroad/where-can-you-study-abroad-free

****