

METRIC COMPARISON BETWEEN GOOGLE SCHOLAR AND RESEARCH GATE FOR RATED SCIENTISTS IN SOUTH AFRICA

Arthur James Swart*

Central University of Technology

drjamesswart@gmail.com



Reception: 21/02/2023 **Acceptance:** 21/04/2023 **Publication:** 10/07/2023

Suggested citation:

Swart., A. J. (2023). **Metric comparison between Google Scholar and Research Gate for rated scientists in South Africa.** 3C Tecnología. Glosas de innovación aplicada a la pyme, 12(2), 15-29. <https://doi.org/10.17993/3ctecno.2023.v12n2e44.15-29>

ABSTRACT

The successful promotion of an academic at an institution of higher learning is affected, to a large degree, by the publication record of the applicant. This record is usually updated in a resume, portfolio, or an online database, such as Google Scholar, Research Gate, or LinkedIn. The purpose of this article is to present a metric comparison between Google Scholar and Research Gate for rated scientists who are employed at universities of technology in South Africa. This may help to establish notable similarities or differences between rated scientists from specific universities and identify which platform they prefer to maintain their publication records. A snapshot quantitative study is used where the total number of citations, h-index values, and Research Gate scores were collected and analyzed. Results indicate that Research Gate has the highest number of authors present from the six universities of technology, with Google Scholar recording the highest number of h-index values for these authors. Only 134 out of the 181 (sample size) authors maintain their records on both databases. It is recommended that researchers in higher education use at least one online database to maintain a record of their publications, thereby enhancing the visibility of the research done at a university and enabling a more valid snapshot of the performance achieved by each researcher.

KEYWORDS

Quantitative, data analytics, citations, h-index

INDEX

ABSTRACT

KEYWORDS

1. INTRODUCTION
2. NRF RATING APPLICATION IN SOUTH AFRICA
3. STUDY CONTEXT
4. QUANTITATIVE RESEARCH METHODOLOGY
5. RESULTS AND DISCUSSIONS
6. CONCLUSIONS

REFERENCES

ABOUT THE AUTHOR

1. INTRODUCTION

“It takes 20 years to build a reputation and five minutes to ruin it. If you think about that, you'll do things differently” [3]. Warren Buffet, a successful American businessman, uttered these words that indicate that it takes time to build a good reputation. For researchers and academics in Higher Education, these words ring true. They need time to complete publications, successfully supervise postgraduate students, contribute to curriculum development, provide evidence of good teaching, serve on various committees, and engage in community development as part of their key performance agreement. This usually contributes to career development, as they progress through the ranks to a full professor. Being a full professor means not only having a wealth of experience but also status and institutional power that can be put to positive use [9]. However, the progression from a junior lecturer to a full professor takes time (12 years or more) in which individuals build a reputation that identifies them among the leaders in their fields of expertise who constantly produce high-quality research outputs.

This reputation is what the National Research Foundation (NRF) in South Africa (SA) seeks to evaluate. The rating of individuals is based primarily on the quality and impact of their research outputs over the past eight years, taking into consideration the evaluation made by local and international peers [13]. This evaluation aids the NRF in identifying an appropriate category into which researchers may be placed (Categories A, B, C, P and Y). Category Y is for promising young researchers while Category A is for leading international researchers. They are called “rated researchers” in SA, with the term “rated scientist” also applicable as used in this article. Their research outputs may include peer-reviewed journal articles, conference papers, research monographs, and patents. An important requirement for applicants is the inclusion of their h-index values from three different databases, namely Web of Science, Scopus and Google Scholar.

These databases usually include a list (which may not be complete) of the research outputs of a researcher and may also serve as a repository for some of them. For example, Google Scholar (GS) is a searchable database that allows users to access information, cross reference that information with other sources and keep up with new research as it is released. It also allows users to access journal articles, conference papers, academic books, pre-prints, theses and dissertations, abstracts and other scholarly literature [4]. Research Gate (RG) is a similar database that places more emphasis on social networking among researchers. It allows users to share their research publications, find collaborators, access job boards and ask and answer questions in real-time [19]. The metrics (citations, h-index, scores) of these databases should not be considered as the pinnacle of one's reputation, but should rather be considered as an initial stepping-stone to establishing the reputation [17]. Although both databases may provide similar publication ranks [18], great differences have been found between the number of citations present on GS and RG [14]. This is partly due to which database is regularly maintained by the researcher, who often has to

manually add his or her latest outputs. The following research questions are therefore posed:

1. What percentage of NRF-rated scientists from universities of technology in South Africa use either Google Scholar or Research Gate to maintain their research publication record?
2. What correlation exists between the annual output-to-input ratio of students to the number of rated scientists present at the six universities in South Africa?
3. What correlation exists between the h-index values present on Google Scholar and Research Gate for specific NRF-rated scientists?
4. What correlation exists between multiple and single authorship for specific NRF-rated scientists when considering their highest citation count for these publications?

The purpose of this article is therefore to present a metric comparison between GS and RG for rated scientists who are employed at universities of technology in SA. This may help to establish notable similarities or differences between rated scientists from specific universities and identify which platform they prefer to maintain their publication records. A quantitative study is used where the total number of citations, h-index values and RG scores were collected and analyzed. A snapshot quantitative study is used to gather and analyze this data from both databases. The article starts with a brief discussion of the rating application submitted to the NRF, followed by the context of the study. Results, discussions and conclusions follow.

2. NRF RATING APPLICATION IN SOUTH AFRICA

The NRF of SA was established on 1 April 1999 with the mandate to contribute to the national development of research and human capacity, national facilities, national science system and science engagement. The NRF rating system is a key driver to build a globally competitive science system in SA. It is a valuable tool for benchmarking the quality of researchers against the best in the world where the rating categories are allocated based on a researcher's recent research outputs (primarily publications from the past eight years) and impact as perceived by international reviewers [13].

Researchers from institutions of higher learning and science councils are invited to apply, with the submission deadline usually around 15 February of each year. The outcome of the rating process is usually communicated at the end of the calendar year. Unsuccessful applications require applicants to wait three years before re-applying. Successful applicants are required to re-apply every 5 years, thereby providing them with the opportunity to improve their rating towards A Category researcher.

A predefined template needs to be populated, which includes a detailed resume and publication list. Four key critical sections of the application are discussed in the paragraphs below. Possible reviewers must also be suggested by the applicant, where the majority should be international peers.

The section called “The Best Research Outputs in Last 8 years” requires the applicant to list his or her top 5 research outputs for the past 8 years, substantiating each one. This motivation should feature quantitative results such as the top-ranked conference in one’s field where a paper was presented, the number of current citations for a book chapter, the impact factor of the journal that published an article and which universities or institutions are making use of an output in a specific way. If the research output was multi-authored, then an explanation of the applicant’s contribution is required.

The applicant must highlight his or her niche area of research in the section called “Brief Description of Completed Research”. This requires the applicant to focus in on a specific field of research, thereby enabling the acquisition of expertise within it. Having multiple focus areas does not contribute to a successful application. A chronological discussion of the progress made over the past 8 years is required, with the niche area featured in each paragraph. Specific achievements for each year should be listed in separate paragraphs, with quantitative results or achievements taking precedence.

The “Self-assessment of Research Outputs” section involves a brief personal discussion of the top 5 selected research outputs for the past 8 years. In this section, the applicant needs to indicate how the outputs have contributed to the niche area of the researcher and what personal development (e.g., new knowledge or skill) occurred in each case.

The final section entitled “Ongoing and Planned Future Research” calls on the applicant to again mention his or her niche area. The importance of this area should be noted here, and what further contributions to it are envisioned by the applicant. This section would typically become the “Brief Description of Completed Research” in 5 years when the applicant re-applies to the NRF for the rating (only valid if the initial application is successful).

Several key benefits accrue to rated scientists in SA [17]. Firstly, they receive a once-off financial grant from the NRF that must be used towards their future research. Secondly, they may receive an additional financial incentive from their institution. Thirdly, they become more valuable and marketable within the sphere of higher education. Fourthly, they improve their research reputation among their peers and colleagues. Fifthly, university students can benefit from exposure to rated scientists, who are usually motivated, dedicated and well-informed in their field of study.

3. STUDY CONTEXT

The context of this study is limited to six universities of technology in SA that seek to improve the number of NRF-rated scientists at their institutions (see Table 1). Two key benefits are derived from this, including additional research funding from the NRF and greater awareness of the research ability and capacity of academics at an institution.

Table 1. Data from four online databases for the six universities

Acronym	Full university name	Enrolments in 2020	Graduations in 2020	Ratio	No. of NRF Scientists	Scientist-to-student ratio
CPUT	Cape Peninsula University of Technology	32383	7593	0,23	58	558
CUT	Central University of Technology	21103	4622	0,22	19	1111
DUT	Durban University of Technology	33933	10034	0,30	41	828
MUT	Mangosuthu University of Technology	14652	2805	0,19	6	2442
TUT	Tshwane University of Technology	62482	14897	0,24	47	1329
VUT	Vaal University of Technology	20293	3566	0,18	10	2029

This table lists the acronym of the university that will be used in this article along with the number of enrolments and graduations for 2020, as indicated on the NRF website. The Department of Higher Education and Training (DHET) in SA provides the NRF with aggregated staff and student data collected through its HEMIS (Higher Education Management Information Systems) database [11].

These statistics help to provide an output-to-input ratio of students (which may also be termed the efficiency of a system by dividing the output by the input and multiplying by 100%) that may be correlated to the number of rated scientists present at the institutions. This can help to determine if a statistically significant relationship exists between these two variables which could indicate that a higher number of rated scientists may lead to a higher output-to-input ratio for an institution. The number of enrolments may also be divided by the number of rated scientists to obtain a scientist-to-student ratio. The expression lecturer-to-student ratio is more common in the literature and can be affected by large classes, inadequate infrastructural facilities and poor management support. One study from Nigeria listed a very poor lecturer-to-student ratio of 1:800 [6], which can impact the quality of teaching and student academic success. These last two factors may also be influenced by the scientist-to-student ratio at an institution, as scientists are motivated, dedicated and experts within their profession, who can easily respond to related questions from their students. Wood and Su [20] emphasize that an excellent lecturer is someone who is committed and dedicated to their profession while Mardiana [8] states that lecturers must have a teaching competency certificate related to performance. Subsequently, an excellent

lecturer may be defined as one who holds a teaching qualification and who is regarded by his or her peers as an expert in a specific field of study. A rated scientist fulfills 50% of this definition, providing students with a wealth of expertise to draw from, thereby improving their professional knowledge.

4. QUANTITATIVE RESEARCH METHODOLOGY

A snapshot quantitative study is used where the total number of citations, h-index values and RG scores were collected and analyzed. A snapshot is used to create a static image of dynamic content that may be used as a benchmark for future research. This was done in the month of April 2022. Present-day data would not correlate to the results presented in this article, as this type of data is dynamic with new citations and publications regularly being recorded. MS EXCEL was used to manually capture and process the data. This included the total amount of citations per author, the highest citation count for a sole and multi-author publication, and the h-index value from the GS database. From the RG database, the RG score, h-index values and h-index values excluding self-citations were captured. The author names were obtained from the NRF website which provides a downloaded EXCEL sheet of all rated scientists (researchers) in SA [12]. The search feature available in GS and RG was used to obtain the required data for each researcher.

Before the static images could be generated, the data had to be processed. A table listing all six universities of technologies with the number of rated scientists per institution was first generated. This table also included the number of student graduations and enrolments for the six universities for the year 2020. The number of rated scientists per category were then determined along with a scientist-to-student ratio. GS citations and RG scores were processed for each university (maximum, mean and median values being calculated). The maximum value represents the highest total citation count for a single researcher while the median helps to evaluate the contribution to the research culture of the institution by all the researchers as a collective.

The h-index values between GS and RG were then compared using the maximum and median values along with a Pearson correlation value. This correlation required that researchers be present on both databases which equated to 134 out of the 181 rated scientists (total sample size). The h-index from RG was further analyzed by contrasting the difference between including self-citations to the exclusion thereof. Finally, a comparison between the highest citation count for a sole and multiple-author publication was processed using data from GS.

5. RESULTS AND DISCUSSIONS

Figure 1 presents the number of rated scientists listed on the NRF website according to five different categories of registration. No researchers from the six universities are currently rated in Categories A and P. This indicates that these

researchers are still progressing toward being recognized as leading international researchers within their field of study. It also indicates that there are no researchers younger than 35 years of age (Category P).

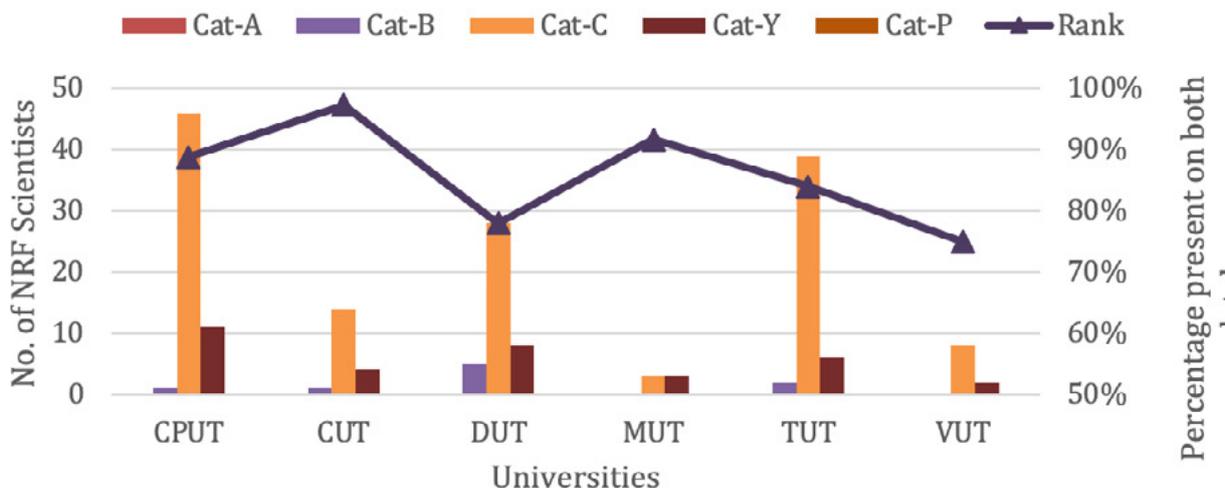


Figure 1. NRF Scientist categories for the six universities along with a ranking percentage of scientists present on both platforms (average value between GS and RG)

The dominant category is C, with a total number of 138 out of the possible 181 rated scientists. This category is for established researchers with a sustained recent record of productivity in their field who are recognized by their peers as having produced a body of quality work (being coherent and ongoing) and have demonstrated the ability to conceptualize problems and apply research methods to investigate them [13]. Noteworthy is the percentage of rated scientists who are present on GS and RG. CUT ranks first in this regard, with 97% of their researchers present on both databases. The lowest presence (75%) is currently found with VUT. High-profile publications of scholars may provide a glimpse of the quality and relevance of an institution’s research regarding global trends [16], which may be enhanced by using more than one online database. Figures 2 and 3 show the results of the GS citations and RG scores.

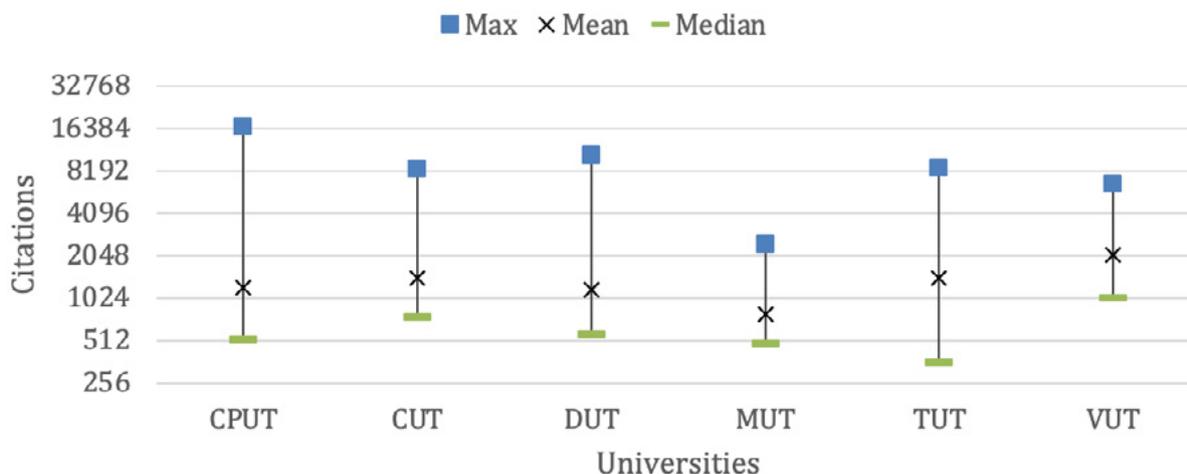


Figure 2. Google Scholar citations

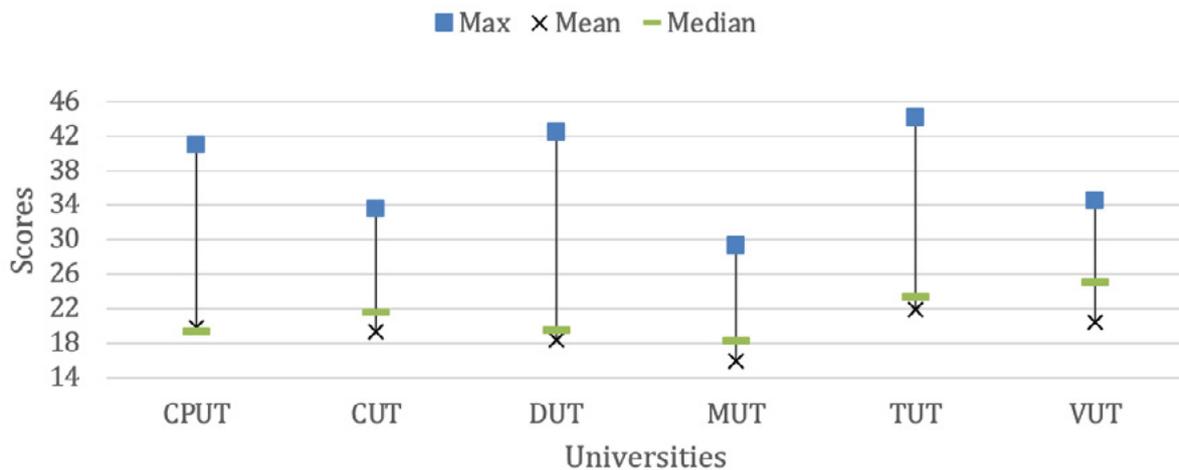


Figure 3. Research Gate scores

From Figure 2 it is evident that individual researchers from four out of the six universities have produced several high-profile publications that have resulted in a maximum total citation count of more than 8192 times (see the Max value). MUT has a researcher with the lowest maximum total citation count (just over 2048) when considering all the universities. This may be correlated to Table 1, which shows that MUT also has the lowest number of rated scientists with the highest scientist-to-student ratio. A possible reason for this low maximum total citation count may be that the university is still trying to establish a good research culture among its academics. Developing positive, supportive, and productive research cultures is a complex and progressively important component of institutional and academic lives [7]. The establishment of increasing productivity and quality metrics, as well as frameworks for measuring impact and research environment, means that the important role of a good research culture should not be ignored. The low median value for TUT is ascribed to the fact that 11 of their rated scientists do not maintain a profile on GS. However, the median is consistently smaller than the mean thereby indicating a positive skew.

A similar profile is noted in Figure 3, where MUT again has a researcher with the lowest maximum RG score of 29. This suggests a similarity between the two databases. However, where GS registered the highest total citation count with CPUT, RG registers its highest score with TUT. Noteworthy also is the gap differences between the mean and median scores between the two databases. GS consistently shows a higher mean-to-median value for its citations, while this is reversed for the RG scores. One must bear in mind that the RG score is calculated using several variables, whereas the GS citation count is purely based on the total of citations to all the published work of a researcher. It has been noted that the RG score should not be used to compare institutions on research quality [5], as is evident by the negatively skewed distribution of data in Figure 3 (median greater than the mean). This infers that the distribution of scores is not symmetrical and there are more extreme scores in the bottom 50% than in the top 50%. Figure 4 considers the h-index.

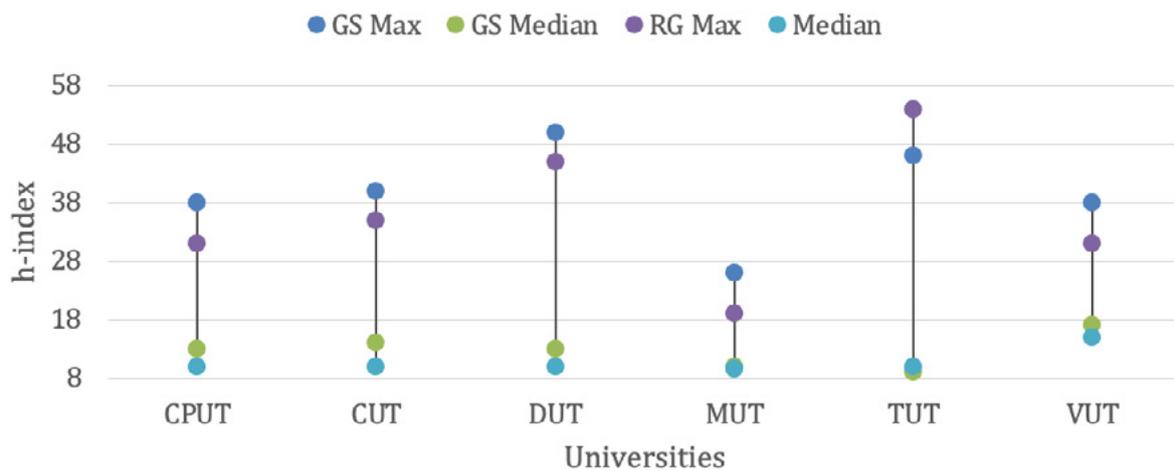


Figure 4. Google Scholar (GS) to Research Gate (RG) h-index comparison

A comparison between the h-index values of GS and RG is done using the highest h-index for a single researcher (see GS Max) and the median for all the researchers as a collective. In this case, five out of the six universities have the potential to gain at least one Category A NRF researcher in the future, as their h-index lies above 38 according to GS. RG paints a similar picture. The question of what number to categorize as a good h-index is controversial [1]. However, it is generally accepted that a h-index of 20 over 20 years of scientific work is considered good while a score of >40 in the same career duration would be outstanding. A statistically significant relationship ($p = 0,893$) was found between the h-index values present on GS and RG for 134 scientists present in both databases. All median values are above 8, which suggests that at least 50% of all rated scientists at the six universities are contributing to the research culture of the institution.

Figure 5 illustrates the results of the h-index values from GS when considering self-citations. The maximum number (Max) represents a decrease in the h-index for a single researcher if his or her publications that include self-citations are excluded from the analysis. In this case, the maximum is 8 from CUT. It is encouraging to see that all six universities have a low median value of 0 or 1, thereby indicating that at least 50% of their rated scientists do not overly engage in the use of self-citations. Self-citations can be a means for researchers to place their ongoing research in the context of their prior research [15] or they can be manipulated by a researcher to affect their influence and recognition [2].

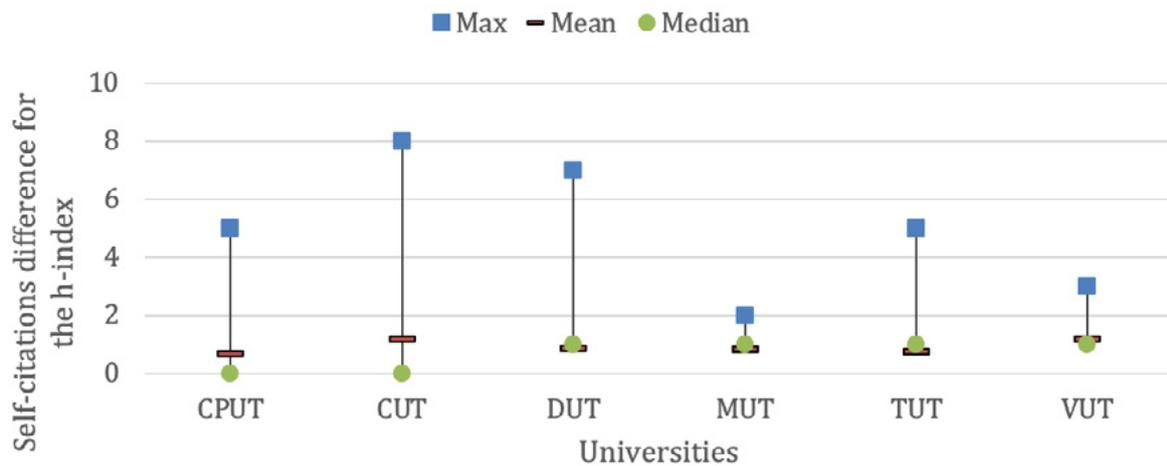


Figure 5. Google Scholar h-index difference when considering self-citations

Figure 6 portrays a comparison between the highest citation count for a sole-author publication versus a multi-author publication for all 181-rated scientists. Consider two examples highlighted with enlarged black circles. The first example is of a researcher who has a sole-author publication with 8 citations and a multi-author publication with 3 citations. The second example shows a researcher with a sole-author paper with 1 citation and a multi-authored paper with 4096 citations. Recall that these values represent the highest citation count for publications from individual researchers. The first example could suggest that the researcher is balancing their research between multi-author (collaborating with fellow researchers) and sole-author (working independently) publications. The second example suggests that this researcher is primarily producing multi-authored publications (engaging only in collaborative work). Ideally, a researcher should be balancing sole and multi-author publications that result in a citation count for both that reside within the box that is shown in the figure. This is currently not the case, as no statistically significant relationship ($p = 0,148$) was found between the highest citation count for a single-author and multi-author publication.

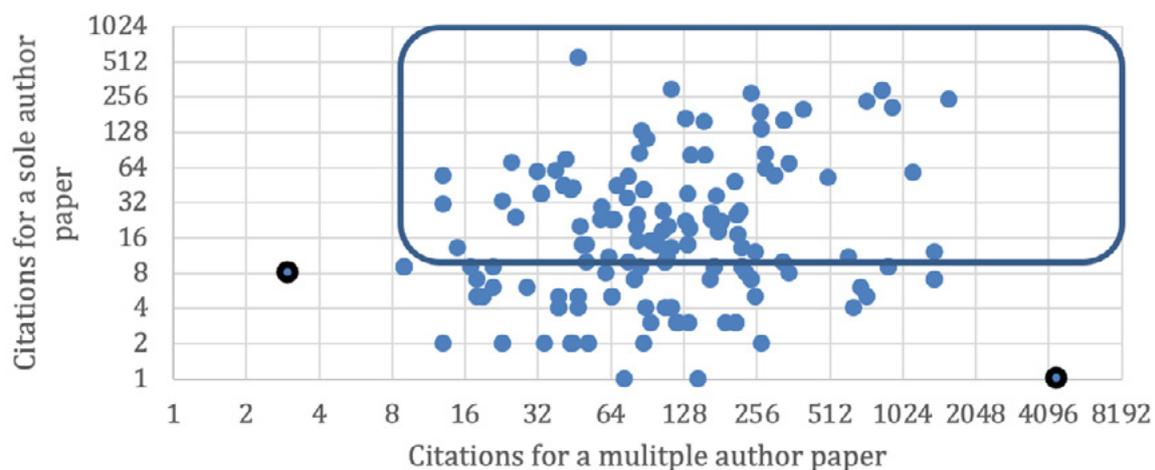


Figure 6. Google Scholar comparison for the highest citation count for a multiple-author publication versus a sole-author publication

It must be noted that a researcher needs to demonstrate the ability to publish on his or her own. This demonstrates the ability of the researcher to conceptualize, synthesize and analyze a research publication. Engaging solely in collaborative work does not convey this demonstration, as co-authors may have conceptualized the publication or even gathered all the data. Being a sole author does represent a significant responsibility, as it means that one has full autonomy and accountability for producing papers worthy of publication. It also enables one to avoid engaging in the somewhat contested space of author order negotiation [10]. It also removes the debate of what contribution each author made to the publication as a percentage that needs to be described in the NRF rating applications.

6. CONCLUSIONS

The purpose of this article was to present a metric comparison between GS and RG for rated scientists who are employed at universities of technology in SA. This helped to establish several notable similarities and differences between rated scientists from specific universities and helped identify which platform they prefer to maintain their publication records. Answers to the original four questions posed under the introduction are now presented.

1. What percentage of NRF-rated scientists from universities of technology in South Africa use either Google Scholar or Research Gate to maintain their research publication record?

From Figure 1, it was shown that CUT ranks first in this regard, with 97% of their researchers present on both databases. A further breakdown of the percentages shown in this figure indicates that 85% of the 181-rated scientists maintain a GS profile, while 87% of them maintain an RG profile. It was further found that the citations from GS provide a positively skewed distribution while the RG scores indicate a negatively skewed distribution (see Figures 2 and 3). This suggests that at least 50% of the rated scientists are achieving a higher citation count while their RG score remains low. However, maintaining at least one profile may improve international awareness of the quality and relevance of an institution's research regarding global trends.

2. What correlation exists between the annual output-to-input ratio of students to the number of rated scientists present at the six universities in South Africa?

Table 1 indicated that DUT achieved a superior output-to-input ratio in 2020 of 0.30 while having the lowest scientist-to-student ratio of 828. A statistically significant relationship ($p = 0,694$) was found between the number of rated scientists and the output-to-input ratio of students at the universities. This suggests that student academic success can be influenced by a higher number of rated scientists at an institution, as these scientists are usually motivated, dedicated and experts within their profession, who can easily respond to related questions from their students.

3. What correlation exists between the h-index values present on Google Scholar and Research Gate for specific NRF-rated scientists?

From Figure 4 it was shown that the h-index values from both databases follow a similar trend. A statistically significant relationship ($p = 0,893$) was found between the h-index values present on GS and RG for 134 scientists present in both databases. This indicates a measure of reliability where rated scientists can make use of either database to maintain their publication record.

4. What correlation exists between multiple and single authorship for specific NRF-rated scientists when considering their highest citation count for these publications?

No statistically significant relationship ($p = 0,148$) was found between the highest citation count for a single-author and multi-author publication for the 181-rated scientists. Ideally, a researcher should be balancing sole and multi-author publications that can enable one to demonstrate independent learning and collaborative work, which is currently not the case as seen in Figure 6. Figure 5 further indicated that at least 50% of the rated scientists from all six universities do not overly engage in the use of self-citations. This would lend credibility to the data presented in this article, as well as to the databases, as the majority of the citations have not been self-inflated by the researchers but represents the active interest of peers in the quality of their work.

A limitation of this study relates to the use of data from only six universities of technology in SA. However, the results of this study can serve as a benchmark for future research-related metric comparisons between GS and RG. The metrics on these two databases should not be considered as the pinnacle of one's research achievement or reputation but should rather be considered as an initial stepping-stone to establishing the reputation. It is further recommended that researchers try to find a balance between multi-authored and sole-authored publications, as this can simply be the application process for rating where different author contributions need to be explained.

Furthermore, it is recommended that management at universities mandate that their academic staff maintain their publication record on at least one of the many available databases in this regard. This can enhance the visibility of the research done at the university and enable an easier review of the achievements of staff for personal performance management purposes and promotion.

REFERENCES

- (1) Ali, M. J. (2021). Forewarned Is Forearmed: The h-Index as a Scientometric. *Seminars in Ophthalmology*, 36(1-2), 1-1.
- (2) Bayram, A. (2021). Author Self-Citation in the Turkish Otorhinolaryngology Literature. *Turkish Archives of Otorhinolaryngology*, 59(3), 210.

- (3) Brainy Quote. (2020). Homepage. Retrieved from <http://www.brainyquote.com/quotes/>
- (4) Delfino, D. (2019). 'What is Google Scholar?': What you need to know about Google's database for students, researchers, and other curious minds. Business Insider. Retrieved from <https://www.businessinsider.com/what-is-google-scholar?IR=T>
- (5) Iatsyshyn, A. V., Popov, O., Kovach, V., Iatsyshyn, A. V., Artemchuk, V., Radchenko, O., ... Kovalenko, V. (2021). Formation of the scientist image in modern conditions of digital society transformation. Paper presented at the Journal of Physics: Conference Series.
- (6) Ifedili, C. J., & Ofoegbu, F. (2011). Managing entrepreneurship education in Nigerian universities. *European Journal of Educational Studies*, 3(1), 101-109.
- (7) Khoo, T. (2021). Creating spaces to develop research culture. *International Journal for Academic Development*, 1, 1-13.
- (8) Mardiana, H. (2021). Lecturers in Adopting Digital Literacy towards Innovation Technological Change. *Zien Journal of Social Sciences and Humanities*, 1(1), 36-48.
- (9) McDaniel, K. N. (2019). You're a Full Professor. Now What? *The Chronicle of Higher Education*, 65, 21.
- (10) Merga, M. K. (2015). Thesis by publication in education: An autoethnographic perspective for educational researchers. *Issues in Educational Research*, 25(3), 291-308.
- (11) National Research Foundation. (2022a). HEMIS Students. Retrieved from <https://www.nrf.ac.za/information-portal/hemis-students/>
- (12) National Research Foundation. (2022b). NRF Rated Researchers. Retrieved from <https://www.nrf.ac.za/information-portal/nrf-rated-researchers/>
- (13) National Research Foundation. (2022c). NRF Rating. Retrieved from <https://www.nrf.ac.za/rating/>
- (14) Orduna-Malea, E., & Delgado López-Cózar, E. (2017). Performance behavior patterns in author-level metrics: a disciplinary comparison of Google Scholar Citations, ResearchGate, and ImpactStory. *Frontiers in Research Metrics and Analytics*, 2, 14.
- (15) Sandnes, F. E. (2020). A simple back-of-the-envelope test for self-citations using Google Scholar author profiles. *Scientometrics*, 124(2), 1685-1689.
- (16) Swart, A. J. (2018). An Analysis of Master Dissertations: A Case Study of Central University of Technology, South Africa. *AJLAIS, African Journal of Library, Archives and Information Sciences*, 28(2), 211-223.
- (17) Swart, A. J. (2022). Metric Comparison Between Google Scholar And Research Gate For Engineering Academics. *Webology*, 19(3), 1025-1036.
- (18) Thelwall, M., & Kousha, K. (2017). ResearchGate versus Google Scholar: Which finds more early citations? *Scientometrics*, 112(2), 1125-1131.
- (19) University of Melbourne. (2019). Researcher Profiles, Identifiers and Social Networks: Maximise your Impact. Retrieved from https://unimelb.libguides.com/researcher_profiles/rg

- (20) Wood, M., & Su, F. (2017). What makes an excellent lecturer? Academics' perspectives on the discourse of 'teaching excellence' in higher education. *Teaching in Higher Education*, 22(4), 451-466.

ABOUT THE AUTHOR



James Swart received his DTech: Electrical: Engineering degree in 2011 from the Vaal University of Technology. His research interests include engineering education development and energy monitoring of PV modules. He is currently an Associate Professor at the Central University of Technology.