Can student evaluation of a course and teacher be of benefit to the evaluating cohort: timing of evaluation in a rural-based institution of Higher Education

FN. Fon[1*] and M. Sibanda[1]

[1]Department of Agriculture, University of Zululand, P. Bag X1001, KwaDlangezwa 3886
*Corresponding author: FonF@unizulu.ac.za

Abstract. Lecturers are at the centre of all teaching and learning in Higher Education Institutions irrespective of the country and continent. For years now, student module and teachers' evaluations have come at the end of every semester and the outcomes have been promising, but student performances remain a major debate. The key questions here are, who benefits from the feedback? Current cohort or future cohort? Does cohort vary in both intrinsic and extrinsic characteristics? If timing is restructured, would it benefit or affect student perspective on evaluation? This study investigated the effect of timing (early September (SeptEV) and November (NovEV)) of student course and teacher's evaluation on cohort benefit and evaluation perspectives. A third-year module (small class of 51 students) from the Department of Agriculture in a rural-based University was used. The evaluation was done midway and end of the semester through a questionnaire through Moodle. Feedback from the first questionnaire was used to improve teaching and learning prior second evaluation. The research was single-blinded to avoid bias in the second evaluation. The results indicated that students' perspectives changed positively (P<0.05) towards learning, assessment, and teaching from SeptEV to NovEV. All 10 questions on course evaluation, 3 out of 5 for teacher evaluation, and 5 out of 5 on assessments all improved positively. The biggest improvements were observed in teachers' evaluation; the pace of lecturing (38.2%), communication and explaining clearly (33.5%), and availability for consultation (22%) while for module evaluation multiple application of library resources (20.2%), appropriate technology enhancing learning (20.5%) and workload satisfaction (23.2%) all improved significantly. For assessments, fairness (29.8%), enough preparation time (18.9), timely feedback (18.6%) and relevances to course outcomes (5.2%) all improved (P<0.05) from SeptEV to NovEV. Timely evaluation if carefully designed may not only improve teaching and learning but also benefit the evaluating cohort.

Keywords: Student course and teacher evaluation, higher education, pedagogy.

1. Introduction

Teachers are said to be one of the most important people in every society because they are responsible for knowledge building at the basic education level as well as the development and application of knowledge at Higher education. However, the learners or students are very important because, without them, there will not be any teaching carrier. At Higher Education (HE) institutions, what is taught is dependent on the module outcomes, which are also directly linked to the overall programme outcomes. With the recent drive in the decolonization of the curricula especially in the African continent, it indirectly means that these modules are continuously being modified to interrogate its immediate communities either directly or indirectly. The national development plan of South Africa positions education at the forefront while goal number four of the Sustainable Development Goals of 2030 (SDG2030) emphasized the need for inclusive and equitable quality education that would promote life learning opportunities globally (United Nations, 2015). The Africa Agenda 2063 (AA2063) also emphasized the need for quality education as there is a huge drive for a knowledge-based economy in the African context (African Union, 2015). Both agendas (AA2063 and SDG 2030) have been discussed by different scholars and found overlaps but development for sustainability was at the centre of each (Garfias Royo, Diep, Mulligan, Mukanga, & Parikh, 2022). It is therefore clear that education has been placed as the focal point for any development and higher education institutions have been given the mandate to push this drive (Godonoga & Sporn, 2023). However, the throughput has not been the best in many countries including South Africa (Ndlovu-Gatsheni, 2017). Therefore, there have been many campaigns and workshops brainstorming on possible reasons for such low throughputs. One among the many strategies (Brown, Pather, & Lawack, 2022; Chiwandire & South, 2022) that have been adopted by most, if not all Universities, to try and understand student poor throughput, especially at the higher education, is student module and teacher evaluation (Prieto, Guede-Cid, Cid-Cid, & Leguey, 2023; Rupp, 2023; Skerritt, 2023; Zhang, Liao, Xiong, Zhu, & Wang, 2022). Therefore, a brief scan of how students' modules and teachers' evaluations impact the scholarship of teaching and learning is important.

2. Student module and teacher evaluation dynamics

There is a huge debate about whether student evaluation of the quality of instruction provided in a module as well as teacher evaluation is a true reflection of quality and teachers' performance (Smith, White, Wagner, Kuzyk, & Prera, 2023). This has been argued by many that there are a lot of inconsistencies regarding the use, validity, and consequences in use for student evaluation. This has been demonstrated in cases where both extremely positive and negative scores were assigned to a module instruction and lecturer evaluation by students in the same class (Prieto et al., 2023). However, there is a consensus that the benefits outweigh the disadvantages and will continuously be used in academics by different stakeholders for various purposes (Chen, 2023; Lasekan, 2021; Touran Bahrami et al., 2021). A typical example was given by Smith et al. (2023) where various stakeholders needed it for different purposes, administrators for staff performance, heads of departments to monitor staff activities in the classroom, and faculties for their classes' performances. Some have even argued that peer evaluation stands a better chance for objective evaluation than students. I also strongly, believe that it is subjective as well for the same reasons that flawed student evaluations hence the need for continuous application of student evaluation (M. Husain & Khan, 2016). The next two section separate module and teacher evaluations and briefly highlights their importance in the evaluation of teaching and learning in Higher Education institutions.

3. Module evaluation and benefits

There has been an ongoing debate on who is the best stakeholder to provide feedback module evaluation. Among the different stakeholders (peers, principal, self-evaluation) (Touran Bahrami et al., 2021), students have been chosen by many scholars as the best. While other scholars continue to believe that students' evaluation is more biased (Akins & Murphy, 2019) and preferred peers' evaluation (Akins & Murphy, 2019), M. Husain and Khan (2016) demonstrated that students' evaluation was

more constructive and more objective than those of peers. However, others had divided opinions where students' opinions were confirmed but requested other evaluation methods of evaluation to complement (M. Husain & Khan, 2016). Therefore, besides viewing students as potential clients, their feedback appears to be very important in enhancing module development (Huxham et al., 2017). Module evaluation allows students to have a say on what they think about the module content, methods, outcomes, study materials, grading practices, assessment methods, and criteria (Chae, Kim, Chang, & Chung, 2017). Module evaluation if done properly can have enormous advantages in course development and someone them as listed as follows; i. course design, here the organization, and structure of the course can be reviewed in a way that missing gaps can be filled and more visible by students, ii. Effectiveness and learning experience, the methods of teaching can be reviewed to suit the learning style of a particular cohort as each will have different characteristics, iii. Pedagogical innovations and life-long learning for development, students are often stocked with differences from their diversity and evaluations allow innovating on teaching styles and method foster learning and when successful, it is bound to benefit even future generations, iv. Faculty development and recognition, better module design can benefit other faculty members for collaboration as well as bring personal satisfaction of having some satisfaction for contributing towards course development in the Department or faculty, iv. Student satisfaction and transparency, student get to have a say on how they are assessed and get to comment if the outcomes of the course were met which give rooms for improvement and alignment of content to speak to module outcomes. These assessments are also indirectly linked to the programme outcomes and graduate attributes, v. accreditation and potential ranking, if the process of module evaluation is robust, it can serve as evidence of high standards and can be used for ranking the university for quality. There are potentially many other benefits of evaluating a module but these are just a few that were shared by many researchers (Chen, 2023; Chiwandire & South, 2022; Prieto et al., 2023; Rupp, 2023).

4. Teacher evaluation and benefits

Student evaluation of teaching (SET) as referred to by most scholars in the scholarship of teaching and learning, constitutes two components which are module content evaluation and teacher evaluation. The previous section discussed module content evaluation which differs from teachers' evaluation but somehow still influences teachers' engagement and effective teaching. Such evaluations are very important for both the professional and personal growth of the teacher if the process is robust (Prieto et al., 2023). Student evaluation by students is still a major concern for many scholars as they believe that it is skewed by many factors that can be somewhat biased (Smith et al., 2023). Others strongly believe that peers' evaluation of teachers is more important because of their pedagogical knowledge which can lead to constructive feedback. However, other scholars again seat in the middle because their study showed that students were more objective than their peers, hence the reason why students' evaluation would continuously be at the forefront. Many benefits can come from teachers' evaluation with an overall effect of improving teaching quality and learning experience. Some of the specific advantages of teacher evaluation would include; i. identification of strengths and weaknesses, ii. innovate in teaching strategies to accommodate students with different learning styles, preferences, and needs, iii. improve communication skills to accommodate students from different backgrounds, iv. Innovate in teaching styles that promote student participation in classroom interactive and engaging activities, v. exploration of multiple teaching resources such as technology to foster learning, vi. Assisting lecturer to pick up differences between different cohorts of students and restructuring of teaching strategies, vii. Can motivate lecturers in terms of good practices to continue to innovate in their teaching, viii. Promotion of accountability by the lecturer is key because some lecturers take the teaching practice for granted, ix. Staff development programmes can be informed based on robust teacher evaluations (Touran Bahrami et al., 2021), x. lecturer evaluation can also lead to rewards in terms of performance management systems. However, whether module or teacher evaluation, many factors can skew students' evaluation and may require strong consideration in their design and this will be discussed in the next paragraph.

There is a consensus that lecturers with better publication records in their field are better teachers. There is a very strong debate about this point because it can not be

generalized and publication record is different from teaching. It has also been reported that the gender and looks of the lecturer be it female or male have some bias in their rating. Timing of evaluation is also another factor, especially after a test or at the end of the semester. Reports have shown that students who fail or have low marks tend to score the lecturers low while good students will score higher marks. Anonymity is another factor that must be considered because some students are worried about giving objective feedback when they understand that they will be identified. However, some lecturers do not share the opinion of anonymity as they think that it allows some students to settle scores or use it as a punitive measure which defeats the objective of the evaluation. Finally, some students do not understand the need for the evaluation while others assume that their feedback will never be read hence that objectivity is flawed and would require careful consolidation during evaluation design.

While scanning through the literature more than 95% of all the reviews and research on SET were performed at the end of the semester (Chen, 2023; Cunningham, Laundon, Cathcart, Bashar, & Nayak, 2023). This time of the year is one of the busiest times for any student's academic career. This is because they will either be involved with tests or reading for their exams. This implies that there would be little or no time for the evaluation even when issued face-to-face because their priority is the upcoming tests or exams. If forced, to respond, the objectivity may just be defeated because their response may be biased and possibly lacks honesty and reliability. Some researchers have tried to control this by issuing the questionnaire within class time but it will be flawed because the teacher should not be around. Therefore, inviting guests or peers to oversee the evaluation is key but could reduce students' seriousness. Others have attempted to mitigate this by involving students in the design and their contribution towards the development of the course and its benefit to future students. Online evaluations were introduced as an alternative to mitigate the problem of timing and this was reported by many as a success and convenient to many since students had enough time to do it when they wanted (Smith et al., 2023). Although this method is used by the majority of researchers currently, many students are still not participating. This implies that the issue of time may still be a major problem besides the lack of electronic devices, data, and connectivity that can affect the online process. Two stud-

ies (one review and one experiment) advocated for continuous evaluation within the semester where students would have enough time and without pressure. The review was only suggesting it as a potential solution to mitigate end-of-semester lower participation. The experiment demonstrated that semester evaluations would not only
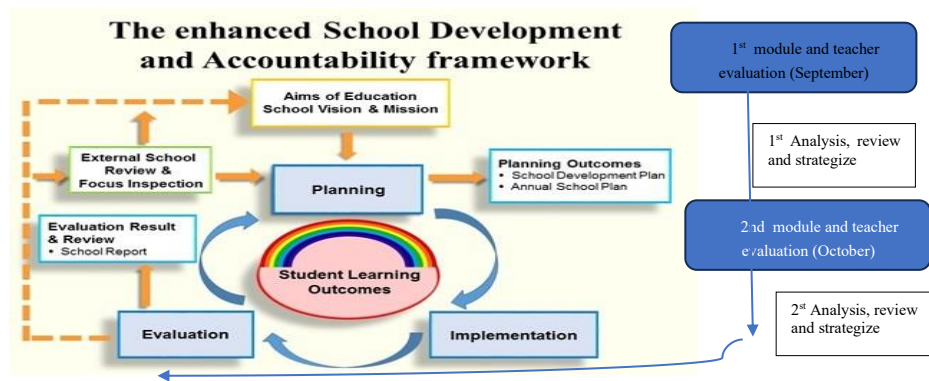

Figure 1: A modified conceptual framework for module and staff evaluation. Adopted from Education Bureau (2022).

increase participation but have the potential to benefit the evaluating cohort more than only the future cohort. However, this evaluation was done across modules taught by the same lecturer in a medical school in the same semester. This can still be flawed because it is the same semester but a different module. Therefore, this study saw an opportunity to investigate the effect of student module and teacher continuous (timing) evaluation on the potential benefit to the evaluating cohorts at a rural-based institution of Higher education. The conceptual framework for module and teacher evaluation (Figure 1) was used adopted from an enhanced school development and accountability framework. The is planning at the beginning of the semester based on module outcome also guided by the programe outcome and the University vision of and mission of its graduates. The planning of teaching and evaluation process, it is then implemented. Evaluation process is done twice a year. After the first evaluation, the data

will be evaluated and shared with lecturer to act upon for improvement for potential benefit of the cohort. The second evaluation at end of semester will be used for improvement and planning for subsequent cohorts.

5. Method

To be more specific, the study investigated the effect of timing (early September (SeptEV) and November (NovEV)) of student module and teacher evaluation on cohort benefit and evaluation perspectives. A third-year module (small class of 51 students) from the Department of Agriculture in a rural-based University was used. The evaluation was done midway and end of the semester through an online questionnaire on the institution's Learning Management System (Moodle). Feedback from the first questionnaire was used to improve teaching and learning prior second evaluation through an online questionnaire. The research was single blinded to avoid bias in the second evaluation. The survey questionnaires were divided into two major groups A (module evaluation) and B (teacher evaluation), where A was further divided into sub-groups A1 (module content evaluation) and A2 (assessment evaluation). Likert scale from 'strongly disagree' to 'strongly agree,' and the scale is commonly converted to a 1-5 scale to average or code for statistical analysis. Anonymity was maintained for each student and were given the freedom to give anonymous codes to the data at their discretion if need be for performance correlation. For statistical analysis, Statistical Package for the Social Sciences (SPSS) was the software used to analyze the data collected. The chi-square test under crosstab an option of descriptive statistics in SPSS was used to analyze the data since all data collected for module and teacher evaluation were all nominal variables. Chi-square is chosen because of its ability to assess the relationship between categorical variables which are non-parametric and provides a formal hypothesis test for independence. The Chi-square association with the nominal data is considered significant when $P<0.05$.

6. Results
The data generated for students' evaluation of module content is represented in Table 1. The results showed that all responses on module evaluation improved ($P<0.05$)

positively from September (SeptEv) to October (OctEv) as demonstrated by the sum of both agreed and strongly agreed. The highest (P<0.05) improvements were reported from responses on Wload, TechEnL, and LibraryIF while the least improvements for CleerP and SICA from SeptEv to OctEv. A few students disagreed with ColaboT, TechEnL, SICA, and MSEM in SpetEv but all changed them to agreed and strongly disagreed in OctEv.

Table 1: Student evaluation of module content within the semester (September and October)

	Module content evaluation timing (%)										
	September					October					
	1	2	3	4	5	1	2	3	4	5	P-value
ClearP	0	2.4	9.8	63.4	24.4	0	3.6	3.6	50	42.9	0.05
ClearIST	0	4.9	19.5	56.1	19.5	0	3.6	3.6	57.1	35.7	0.05
ColaboT	0	9.8	19.5	51.2	19.5	0	0	17.9	46.4	35.7	0.05
Wload	0	7.3	26.8	56.1	9.8	0	3.3	7.1	60.7	28.6	0.05
TechEnL	0	9.8	22	48.8	19.5	0	0	10.7	57.1	32.1	0.05
LibraryIF	4.9	9.8	34.1	36.6	14.6	0	11	17.9	50	21.4	0.05
SMPis	4.9	9.8	34.1	34.1	17.1	0	14	17.9	35.7	32	0.05
SICA	0	0	22	61	17.1	0	0	17.9	46.4	35.7	0.05
OPIS	0	4.9	19.5	56.1	19.5	0	7.1	7.1	42.9	42.9	0.05
MSEM	0	4.9	19.5	46.3	29.3	0	0	14.3	53.6	32.1	0.05

Likert scale (1=strongly disagree, 2=disagree, 3=neutral, 4=agree and 5=strongly agree), ClearP = Module outcome clearly defined, ClearIST = students' module responsibility communicated, ColaboT=content provided the opportunity for collaborative learning, assessment, and participation, Wload = Achievable semester workload, TechEnL=use of appropriate technology for teaching, LibraryIF = content offers the opportunity to exploit library resources, SMPis =Learning materials were organized in smaller, manageable pieces, SICA = scientific instructions for completing task, OPIS = opportunity for independent study, MSEMvc = make use of additional reading materials

Student evaluation of assessment methods varied (P<0.05) from SeptEv to OctEv (Table 2). All the responses reported for assessment (AssesDW, ASSESFAIR, ASESROUT, TimeOk, Tfback) improved (P<0.05) positively from SeptEv to OctEv. The highest improvement in student responses on assessment from SeptEv to OctEv was reported on ASSESFAIR, TimeOk, and Tfback while AssesDW and ASESROUT were the least improvements reported. ASSESFAIR, TimeOk, and Tfback reported the presence of some strongly disagreed and disagreed students

which was missing for AssesDW and ASESROUT but for disagreed in SeptEv. Apart from ASSESFAIR, all the strongly disagreed students changed their responses towards the positive Likert scale for assessment.

Table 2: Student evaluation of module assessment in the semester (September and October)

	Module content evaluation timing (%)										
	September					October					
	1	2	3	4	5	1	2	3	4	5	P-value
AssesDW	0	2.4	17.1	61	19.5	0	0	14.3	35.7	50	0.05
ASSESFAIR	7.3	4.9	39	41.5	7.3	3.6	3.6	14.3	39.3	39.3	0.05
ASESROUT	0	0	19.5	51.2	29.3	0	7.1	7.1	32.1	53.6	0.05
TimeOk	7.3	9.8	26.8	43.9	12.2	0	10.7	14.3	42.9	32.1	0.05
Tfback	4.9	9.8	34.1	34.1	17.2	0	14	17.9	35.9	32.1	0.05

Likert scale (1=strongly disagree, 2=disagree, 3=neutral, 4=agree and 5=strongly agree), AssesDW = use of multiple assessments to assess module, ASSESFAIR = Assessment were fair, ASESROUT = assessment align to module outcome and teaching, TimeOk = enough time for writing, Tfback = timely feedback after every assessment.

Table 3: Student teacher evaluation in the semester (September and October)

	Student teachers' evaluation timing (%)										
	September					October					
	1	2	3	4	5	1	2	3	4	5	P-value
LAHK	0	2.4	2.4	80.5	14.6	0	0	7.1	25	67.9	0.05
LAWAP	0	4.9	19.5	51.2	24.4	0	0	7.1	21.4	71.4	0.05
LMCPace	7.3	17.1	31.7	36.6	7.3	0	0	17.9	50.0	32.1	0.05
LCCEW	2.4	9.8	39	34.1	14.6	0	0	17.9	42.9	39.3	0.05
LCSTIND	0	0	4.9	34.1	61	0	3.6	3.6	60.7	32.1	0.05

Likert scale (1=strongly disagree, 2=disagree, 3=neutral, 4=agree and 5 = strongly agree), LAHK = lecturer is knowledgeable content-wise, LAWAP = lecturer willingness to attend to students problems,/questions, LMCPace = lecturer moves at the correct pace, LCCEW = lecturer ability to explain and communicate module content, LCSTIND = lecturers challenges student development skills.

Teacher evaluation reports showed varied (P<0.05) responses from SeptEv to OctEv (Table 3). Three (LAWAP, LMCPace, LCCEW) out of the five questions reported improvement (P<0.05) in teacher's performance rating from SeptEv to OctEv. The

number of students that decided to stay neutral for teacher evaluation on LAHK and LCSTIND increased slightly from SeptEv to OctEv. All the strongly disagreed and disagreed responses from all the teacher evaluation reports in SeptEv were eliminated in OctEv but for LCSTIND disagreement.

7. Discussion

The main objective of student evaluation of teaching (SET) in higher education is to foster teaching and learning irrespective of the continent (Cunningham et al., 2023). The process of most SET seems to favour future cohorts because of the timing of the evaluations (end of the semester). This study was designed to answer the question of whether timing would affect student module and teacher evaluation responses if the lecturer was allowed to adjust teaching and learning strategies based on the first evaluation report. The results showed that timing had an overall positive impact on student module and teacher evaluations from SeptEv and OctEv. In terms of module content evaluation, the student's responses to the same questions in OctEv improved positively (all strongly agreed and agreed) compared to that of SeptEv. This was evident that the lecturer had the opportunity to modify strategies of teaching and learning based on weaknesses that were identified by students. Weakness in module design is very common but improvements on such weaknesses often benefit future cohorts. The results showed that the slight modification improved the module design and presentation of content which is associated with the sudden change of student response in OctEv. Although the number of students that decided to stay neutral in the SeptEv decreased in OctEv, more than 50% of the questions were reported with about 18% of students that remain neutral. The reason for this was not clear but could be associated with a lack of understanding of the questions (Chen, 2023), not being interested in the whole process (Chae et al., 2017), or do not think that their contributions will be read (Musharraf Husain, Khan, & Research, 2016) as reported by other scholars. Therefore, a future involvement of the students in the design and development of surveys could be a better way to decrease the number of neutral responses. While looking at the overall picture of the module evaluation, SeptEv was able to change the perspective of the students in OctEv as follows, Strongly disagree (0.98 to 0%), disagreed (6.36 to 4.5%), neutral (22.68 to 11.8%), agreed (51 to 50%) and strongly agreed (19

to 33.9%). While the negative attitudes were increasing, the positive ones were increasing. This finding shows that it will be right to confirm that mid-semester evaluations will be beneficial for the current cohort. For the module assessment, a similar trend was observed in terms of overall responses from SeptEv to OctEv as follows; strongly disagree (3.9 to 0.72%), neutral (27.29.68 to 13.6%), agreed (46.3 to 37.2%) and strongly agreed (17.1 to 41.5%) but for disagreed (5.3 to 7.1%). Assessment evaluation among the many studies that have been discussed only showed that emotions are always flowing through when students are responding instead of being objective. Many studies have shown trends of students who are doing well scoring content or teacher with higher scores while does that are failing, score lower marks (Stroebe, 2016). It is one of the reasons that was given to the higher percentage of strongly disagreed and disagreed scores observed in assessment evaluation. Even with the higher scores, there were still positives to take from this study because the overall numbers decreased from SeptEv to OctEv.

Teachers' evaluation by students is still one of the most debated topics in the scholarship of teaching and learning because of its potential biasedness and lack of pedagogical knowledge. However, the majority still agrees that it should be used because some of its feedback has been very useful in identifying strengths and weaknesses, shaping teaching styles, creating an inclusive environment for learning as well as communication skills and many other factors already discussed in the introduction. The overall Likert scores for teachers' evaluation decreased for the lower scores but increased toward the higher scores as follows; strongly disagreed (1.9 to 0%), disagreed (6.8 to 0.7%), neutral (19.5 to 10.7%), agreed (47.3 to 40.0%) and strongly agreed (24.4 to 48.6%). This meant that the SeptEv did motivate the lecturer to improve on their pace, communication skills, and create more time to attend to students' problems and feedback timing. The feedback, low scores also resorted from the time of the SeptEv which had not had any major assessment that required feedback but improved in OctEv because they have hard opportunity to get feedback. The lecturers' knowledge of the subject was not 100% and even decreased in OctEv but this was associated with the relatively higher percentage of neutral students unlike in the SeptEv that they disagreed. It is clear from this study that student evaluation is here to

stay and academics and administrators will continue to use it but the planning, designing, timing, and students involved should all be considered for the robustness of evaluation.

8. Conclusion

The study showed that student module and teacher evaluation twice (September and October) within a semester had beneficial effects on both the current cohort and the lecturer. This was demonstrated by the increased positive responses for module content, assessment, and teacher evaluation from SeptEv to OctEv. The lecturers' exposure to SeptEv evaluation analysis was very instrumental for strategic improvement plans that were responsible for the higher scoring by students. However, further studies are requested for the correlation of students' performances with the response to current questions for a better understanding and mitigation strategies for teaching and learning experiences.

9. Acknowledgment

The study would like to acknowledge the students and staff members who participated in the study and the Department of Agriculture and the University of Zululand for providing the resources for carrying out this research.

10. References

African Union. (2015). *Agenda 2063: the Africa we want - framework document*. In (pp. 1-172). doi: https://doi.org/10.2345/0899-8205-44.1.49

Akins, L. M., & Murphy, L. (2019). Panel: Peer-Based Faculty Evaluation v. Student Evaluation of Teaching. *J Journal of Collective Bargaining in the Academy*.

Brown, V., Pather, S., & Lawack, V. (2022). *A Re-Imagined Institutional Student Retention and Success Framework to Enhance Holistic Student Support.* Paper presented at the Proceedings of The International Conference on Future of Teaching and Education.

Chae, S. J., Kim, M., Chang, K. H., & Chung, Y. S. (2017). Potential bias factors that affect the course evaluation of students in preclinical courses. *Korean J Med Educ, 29*(2), 73-80. doi:10.3946/kjme.2017.54

Chen, K. (2023). Exploration and Practice of Teaching Quality Evaluation for University Teachers: Taking the Comprehensive Evaluation of "Six Dimensions and Three Elements" as an Example. *Journal of Education and Educational Research, 3*(1), 21-24. doi:10.54097/jeer.v3i1.8057

Chiwandire, D. J. S. o. T., & South, L. i. t. (2022). COVID-19 pandemic lockdown impact on parity of participation for students with mental health challenges in higher education. *6*(2), 73-99.

Cunningham, S., Laundon, M., Cathcart, A., Bashar, M. A., & Nayak, R. (2023). First, do no harm: automated detection of abusive comments in student evaluation of teaching surveys. *Assessment & Evaluation in Higher Education, 48*(3), 377-389. doi:10.1080/02602938.2022.2081668

Education Bureau. (2022). An enhanced school Development and accountability framework. *https://www.edb.gov.hk/en/contact-us/index.html*.

Garfias Royo, M., Diep, L., Mulligan, J., Mukanga, P., & Parikh, P. (2022). Linking the UN Sustainable Development Goals and African Agenda 2063: Understanding overlaps and gaps between the global goals and continental priorities for Africa. *World Development Sustainability, 1*, 100010. doi:https://doi.org/10.1016/j.wds.2022.100010

Godonoga, A., & Sporn, B. J. S. i. H. E. (2023). The conceptualisation of socially responsible universities in higher education research: a systematic literature review. *48*(3), 445-459.

Husain, M., & Khan, S. (2016). Students' feedback: An effective tool in teachers' evaluation system. *Int J Appl Basic Med Res, 6*(3), 178-181. doi:10.4103/2229-516x.186969

Husain, M., Khan, S. J. I. J. o. A., & Research, B. M. (2016). Students' feedback: An effective tool in teachers' evaluation system. *6*(3), 178.

Huxham, M., Scoles, J., Green, U., Purves, S., Welsh, Z., & Gray, A. (2017). 'Observation has set in': comparing students and peers as reviewers of

teaching. *Assessment & Evaluation in Higher Education, 42*(6), 887-899. doi:10.1080/02602938.2016.1204594

Lasekan, O. (2021). Using Stronge Teacher Evaluation System to Assess the Effectiveness Level of Mr. Brown as an EFL Teacher in the Mind Your Language TV Show: An Attempt to Validate a Reflective Tool to Train Preservice EFL Teachers. *Front Psychol, 12*, 648760. doi:10.3389/fpsyg.2021.648760

Ndlovu-Gatsheni, S. J. J. K. (2017). The emergence and trajectories of struggles for an'African university': The case of unfinished business of African epistemic decolonisation. *43*(1), 51-77.

Prieto, J., Guede-Cid, R., Cid-Cid, A. I., & Leguey, S. (2023). Major increases in teachers' performance evaluations: Evidence from student evaluation of teaching surveys. *Tuning Journal for Higher Education, 10*(2), 105-125. doi:10.18543/tjhe.2299

Rupp, M. T. (2023). A Critical Look at Student Evaluations of Teaching. *American Journal of Pharmaceutical Education*, 100136. doi:https://doi.org/10.1016/j.ajpe.2023.100136

Skerritt, C. (2023). School autonomy and the surveillance of teachers. *International Journal of Leadership in Education, 26*(4), 553-580. doi:10.1080/13603124.2020.1823486

Smith, B. O., White, D. R., Wagner, J., Kuzyk, P., & Prera, A. (2023). Distributional properties of the statistic of online student evaluations the mean does not mean what you think it means. *Studies in Higher Education*, 1-18. doi:10.1080/03075079.2023.2211079

Stroebe, W. J. P. o. P. S. (2016). Why good teaching evaluations may reward bad teaching: On grade inflation and other unintended consequences of student evaluations. *11*(6), 800-816.

Touran Bahrami, B., Amiri, S., Mahmoodi, Z., Marjan Haj Firooz, A., Sara, R., Hadi Soleymani, S., & Latifi, A. (2021). Comparison of Teacher Evaluation Indices from the Perspective of Students and Teachers at Alborz University of Medical Sciences. *Journal of Research Development in Nursing and Midwifery*. doi:10.52547/jgbfnm.18.2.45

United Nations. (2015). Transforming our world: the 2030 Agenda for sustainable development - A/RES/70/1 United Nations (UN), . *United Nations (UN) General Assembly*. doi:10.1163/157180910x12665776638740

Zhang, Z., Liao, W., Xiong, X., Zhu, R., & Wang, J. (2022). Evaluation of online game-based and traditional teaching methods in cephalometric landmarks identifications. *Am J Orthod Dentofacial Orthop, 161*(6), e534-e543. doi:10.1016/j.ajodo.2021.09.014