Short Research Article

The Quality of Classroom instructions from student's perspectives

Abstract

This research examined students' perceptions of education quality among middle and higher secondary schools in Bhutan. Student evaluation of education quality facilitates data-driven decision making in a number of ways. Based on a cross-sectional survey design, a one-time data was collected from 312 students studying in class 9 to class 12, using student evaluation of education quality (SEEQ), developed by Marsh.SEEQ measures students' perception on learning, teacher's enthusiasm in teaching, lesson organization, group interaction, individual rapport, breadth or coverage, examinations, assignments, and perceptions about teacher effectiveness.Both descriptive and inferential statistics were used to analyse the data. Results indicated that the participants had a favourable view of education quality, and that there were no significant differences when the data was disaggregated by gender, except for two sub-scales group interaction and individual rapport. Girls had more favourable view of teachers creating greater opportunities for group interaction and individual rapport compared to boys.However, when disaggregated by the marks obtained during the last examinations, the majority of the sub-scales revealed significant differences. It is argued that cultural dispositions may have come into effect, when the students evaluated their teachers. Directions for further research are provided.

Key words:

Evaluation, learning, teacher enthusiasm, group interactions, rapport, learning, assessment.

Introduction

Students' evaluations of teaching (SET) are activities in which students evaluate teachers' teaching effectiveness and quality. Most educational institutions employ SET to improve teaching performance [1, 2]. Student ratings are an efficient method of assessing teaching effectiveness since they can provide: 1) formative input to improve teaching quality; 2) evaluation of teaching effectiveness for promotion choices; and 3) information to students concerning course and instructor selection [3]. Student evaluations of teaching effectiveness are now widely used in many educational institutions and are included in their quality assurance system. Educational institutions use this useful input from student ratings to evaluate teachers and observe their teaching techniques.Student evaluations serve two main goals [4]: 1) they provide teachers with ideas to improve their teaching performance and course delivery, and 2) they provide administration with information to make judgements about teachers or courses.

Background

The Bhutan Professional Standards for Teachers (BPST) mandates that teachers should use student feedback to continuously improve and personalize their teaching to suit the diversity of students preferred ways of learning [5]. The collection of anonymous, systematic student feedback on teachers' teaching could facilitate teachers understanding of the needs and expectations of the students besides that the data could also be used formatively to improve the teaching learning experiences for the students.Some researchers believe that students increasingly see themselves as 'customers' purchasing a service and, as a result, are becoming more vocal in communicating their desires and perceptions of 'good' teaching to schools and individual teachers [6, 7]. Universities have grown increasingly market-oriented, since many students aspire to 'get a degree' rather than 'be learners'[8]. Several studies have demonstrated a

desire among teachers to gain favourable feedback from their students [9, 10], putting pressure on teachers to respond to students' desires and preferences. However, teachers face pressure not only from students and a personal desire to earn positive feedback but are also guided by their curriculum and professional pride to provide students with the knowledge they need for their future careers. However, many teachers report not making changes suggested by the students if the suggestions go against their professional judgement [10].

There appears to be a general consensus that student feedback helps to improve courses [3], although the impact is considered small [11, 12], and the feedback is, in many cases are not taken seriously [13]. Kemberet al., [14] saw no improvement in student ratings despite the use of student questionnaires,possibly because the feedback was not used effectively. Several studies have shown that manyteachers have a positive view of student evaluations as a useful tool to improve courses [15, 16, 17],with a particular interest in items regarding their interaction with students [18]. Teachers with a positive attitude towards feedback are alsomore influenced by it in the design of their courses [16, 19] andreceive higher feedback scores [20]. However, otherstudies show a more negative view of the evaluation where teachers do not perceive the feedback asvalid [21] or see it as only marginally valuable [18]. Theteachers' main interests relate to their own professional development, career and course improvement,rather than in responding to feedback from the students [22].Therefore, it is contingent upon the teacher to either use or not to use the student feedback gathered.

Research objectives

While student feedback is mandated by policy documents and that the use of such mechanisms have the potential to improve the teaching and learning processes, evidences relating to

improvement in Bhutanese educational setting is scarce.Although measurements regarding the quality of education are conducted by divisions within the Ministry of Education and Skills Development, such as the Education Monitoring Division, these lack student perspectives. Therefore, the overall objective of the research is to determine students' perceptions of educational quality. Specificallythis research attempted to quantitatively measure students' perceptions of the quality of learning outcomes and the learning processes, and determine significant differences within the sample.

Research questions

The central question guiding this research is as follows:

What are students' perceptions of educational quality in a higher secondary school?

The sub-questions guiding the research are as follows:

1. What are the students' perceptions about the learning outcomes and the learning processes?
2. Are there significant differences in the perceptions about learning outcomes and processes based on students' gender and the marks obtained in the last examination?

Literature review

SET is used as a measure of teaching performance in almost every institution of higher education throughout the world [23].From a policy perspective, SETs are mandated in the Bhutanese schools as a means to collect feedback on teachers' teaching practices, so that adjustments could be made.Students are considered important stakeholders in the process of gathering insight into the quality of teaching [24]. Although, SET was originally intended primarily for formative purposes, such evaluations came into use for faculty personnel decisions in the 1970s. More

recently, SET procedures have been included as a key mechanism in internal quality-assurance processes as a way of demonstrating an institution's performance in accounting and auditing practices.

Student evaluation of teaching serves three purposes: (a) improving teaching quality, (b) providing input for appraisal exercises, and (c) providing evidence for institutional accountability [14]. In most institutions, SET is obviously used for formative purposes (e.g., as feedback for the improvement of teaching) as well as for summative purposes (e.g., mapping teaching competence for administrative decision-making and institutional audit [10]. These dual usages - and the unresolved tension between them [25] - makes the use of SET fragile. On the one hand, teachers are convinced of the value of SET as an instrument for feedback on their teaching [26]. Data obtained from SET help them to improve the quality of their teaching, as they provide teachers with insight into the strengths and weaknesses of their teaching practice, based on student opinions. For this reason, one can assume that many teachers welcome SET results in order to improve their subsequent teaching. On the other hand, it has been argued that the main purpose of SET involves its use as a measure for quality monitoring, administrative policymaking [27], and for determining whether teachers have achieved a required standard in their teaching practice [4]. This justification for using SET in staff appraisals is related to an increasing focus on internal quality assurance and performance management, which have become subject to the demands of consumer satisfaction [6]. Student satisfaction has come to play an important role in managerial decisions,based on such key concepts as accountability, visibility, and transparency [28]. Teacher performance and the quality of teaching could thus be defined as the extent to which student expectations are met, equating student opinions with knowledge. For this reason, many faculty members have been questioning the validity and

reliability of SET [29]. Their concerns are comprehensible and appropriate as SET results can have serious effects on a teacher's professional career.

Despite the controversies over the administrators' use of SET data, student evaluation of education quality is situated in Vygotsky's [51] social development theory. In social development theory, the role of adults in children's learning may potentially lead to zone of proximal development, and student feedback about teaching effectiveness provides an avenue for the teacher to understand students learning needs and preferences. However, in the absence of student feedback data, it is impossible to understand how students perceive teachers' teaching and an in-depth understanding of students' learning needs and preferences cannot be developed. Hence, SET is a measure teachers can use to facilitate better teaching.

Research methodology

The study was guided by the principles of survey research which fall within the quantitative research design. Survey research in its basic terms is defined as a means of gathering data or information using a self-report questionnaire or interviews [30]. A "survey design provides a quantitative or numeric description of trends, attitudes, or opinions of a population by studying a sample of that population. From sample results, the researcher generalizes or draws inferences to the population" [31, p. 135, 32].According to [33, 34, 35] cross-sectional survey design has the advantage of collecting information on a set of variables in a shorter period of time. Also, according to [35] the additional advantage of a survey is that data can be gathered from a large cohort or participants which can constitute a representative sample from the population, referred to as generalizability. Generalizability refers to the act of making inferences about the unobserved based on the observed [36]. They assert that it is widely-acknowledged as a quality standard in quantitative research.

Data was collected from 312 class 9 to class 12 students from higher secondary schools in Trashigang district. The students were asked to respond to a structured questionnaire which was be made available online. To collect the data, Student Evaluation of Education Quality (SEEQ) developed by Marsh (1982), was used.SEEQ is a 29-item Likert scale, which measures 9 dimensions of education quality, scored from *Strongly Disagree (1)* to *Strongly Agree (5).* The first dimension or sub-scale is learning, which was measured using 4 Likert-items, consisting of statements like *"I have found the course intellectually challenging and stimulating"*. The second sub-scale is enthusiasm, which was measured using 4-items, such as *"The subject teacher was enthusiastic about teaching the course".*Organization was also measured using 4-Likert items, which included statements like "*Course materials were well prepared and carefully explained".* Group interactions was measured through 4-Likert items with items such as "Students were encouraged to participate in class discussions". Breadth was measured with 4-Likert items and example included "*Subject teacher presented the background or origin of ideas/concepts developed in class*". Individual rapport was measured using 4-Likert items such as *Subject teacher made students feel welcome in seeking help/advice in or outside of class".*Examinations was measured using 3-Likert items such as "*Feedback on examinations/graded materials was valuable."* Assignment and teacher quality was measured using two Likert items such as *"Readings, homework, laboratories contributed to appreciation and understanding of subject"* and "*Compared with other instructors I have had, I would say the subject teacher is"* respectively.

Data analysis

To first determine the reliability of the instrument used, Cronbach's alpha for all the subscales were calculated. Cronbach's alpha measures the internal consistency of the items in a sub-scale

and is considered to be one of the reliability measures. According toNunnally and Berstien[37], an alpha of 0.6 and above are considered to be reliable. As shown in table 2, the Cronbach's alpha value obtained for all the sub-scales were greater than 0.6. The reliability determination was followed by both descriptive and inferential statistics.

Results

Participant's demographic characteristics are provided in Table 1. From the sample, 56.8% of the participants were girls. During the last examinations, the majority of the students, corresponding to 41.2% had scored between 50 and 70 percent. In terms of students' interest in the course, 65.3% reported medium levels of interest. In terms of the difficulty of the course, 64.5% of the participants reported that the course difficulty was average.

Table 1

Percentage of participant demographic and important variables

Demographic	*Descriptors*	*%*
Gender (N = 312)	Female	56.8
	Male	43.2
Marks obtained	Below 40%	11.3
	41-55%	36.3
	56-70%	41.2
	71-85%	9.6
	More than 86%	1.6
Level of interest	Very low	1.0
	Low	6.4
	Medium	65.3
	High	24.4
	Very high	2.9
Course difficulty	Very difficult	3.6
	Difficult	26.6
	Average	64.5
	Easy	4.9
	Very easy	.3

The descriptive results of the sub-scales obtained are provided in Table 2. The mean and standard deviation of learning sub-scale (*M* = 3.77, *SD* = .56), indicate that students find the course challenging and intellectually stimulating, and that they are learning the subjects well. Results (*M = 3.86, SD = .68)* also suggest that the subject teachers were enthusiastic about teaching their respective subjects and were able to hook and sustain students' attention in classroom teaching. The results of the organization sub-scale (*M = 3.76, SD = .65),* suggests that the teacher organized their lessons well and used a diversity of strategies to present the subject matter to the students. Group interactions, sharing of ideas, and participation in the classroom was encouraged by the teachers (*M = 3.94, SD = .76).* Students also reported that their teachers were considerate, approachable, and had a good rapport with the teachers (*M = 4.02, SD = .63).* From among the other sub-scales, individual rapport was found to have the greatest mean. The mean and standard deviation also suggests that the teacher adequately covers the breadth of the topics taught in the class (*M = 3.79, SD = .59).* For example, participants reported that the teachers compared and contrasted the implications of various theories taught, and provided alternative and contrasting points of view during lessons. Examinations and assessments were conducted conforming to the policies, rules, and regulations. The overall results suggest that students had a favourable view of the course, teachers' instructional practices, examinations and assignments. This further suggests that students had a favourable view of the education quality that they receive.

Table 2

Cronbach's alpha, mean, and standard deviation of SEEQ sub-scales

N = 312	*n*	*Alpha*	*Mean*	*Std. Dev*
Learning	4	0.65	3.77	0.56
Enthusiasm	4	0.78	3.86	0.68

Organization	4	0.80	3.76	0.65
Group Interaction	4	0.88	3.94	0.76
Individual rapport	4	0.80	4.02	0.63
Breadth	4	0.84	3.79	0.59
Examinations	3	0.75	3.84	0.56
Assignment	2	0.70	3.93	0.65
Teacher	2	0.86	3.67	0.61

The correlation statistics obtained between the sub-scales, provided in Table 3, indicate that the sub-scales were significantly correlated to one another. The correlation coefficients obtained suggested that the correlations were positive, and ranged from moderate to high correlation (Cohen et al., 2018). This indicates that the sub-scales are related to one another and that it effectively contributed to the measurement of education quality.

Table 3

Correlation between the sub-scales in SEEQ

	Learning	*Enthusiasm*	*Organization*	*Group interaction*	*Individual rapport*	*Breadth*	*Exams*	*Assignments*	*Teacher*
Learning	1								
Enthusiasm	.567**	1							
Organization	.607**	.726**	1						
Group interaction	.533**	.656**	.700**	1					
Individual rapport	.462**	.585**	.612**	.659**	1				
Breadth	.595**	.651**	.638**	.593**	.636**	1			
Exams	.448**	.493**	.545**	.551**	.592**	.567**	1		
Assignments	.518**	.496**	.554**	.541**	.470**	.555**	.557**	1	
Teacher	.457**	.620**	.618**	.541**	.502**	.484**	.469**	.404**	1

**. Correlation is significant at the 0.01 level (2-tailed).

Analysis of variance (ANOVA) was conducted to determine if there were significant differences in the participant's scores based on gender. As shown in Table 4, there were significant differences between the scores of male students and female students, especially in group interaction and individual rapport.

Table 4

ANOVA results based on gender

	Female		*Male*			
	M	*SD*	*M*	*SD*	*f(1, 308)*	*Sig*
Learning	3.80	0.53	3.74	0.60	1.07	0.30
Enthusiasm	3.91	0.63	3.80	0.73	2.20	0.14
Organization	3.81	0.57	3.70	0.75	2.30	0.13
Group interaction	4.01	0.66	3.84	0.87	3.78	0.05*
Individual rapport	4.09	0.62	3.93	0.65	4.79	0.03*
Breadth	3.83	0.57	3.75	0.62	1.45	0.23
Exams	3.86	0.56	3.82	0.57	0.27	0.61
Assignments	3.98	0.61	3.87	0.70	2.20	0.14
Teacher	3.70	0.58	3.64	0.63	0.89	0.35

As shown in Table 4, girls perceived that their teachers provided greater opportunities for group interaction (*M =4.01, SD = .66)* compared to male students *(M= 3.84, SD = .65*), and that the relationship was significant at *f* (1, 308) = 3.78, *p =.05.* In a similar manner, girls' participants also perceived that they had better individual rapport with their teacher (*M*= 4.09, *SD* = .62) compared to the male students (*M* = 3.93, *SD* = .65). This difference was also significant at *f*(1, 308) = 4.79, *p* = .03.

ANOVA was also conducted to determine if there were significant differences in the participants' scores based on the marks obtained in the last examinations. To create two groups of participants, the students who scored less than 55% in the last examinations was grouped into one category and the students who scored greater than 55% were grouped into another category. The results obtained are provided in Table 5. Significant differences were obtained for all the sub-scales of SEEQ except for breadth. The results indicate that students who scored more than 55% in the last examinations found the course challenging but yet demonstrated greater learning (*M*= 3.88, *SD*=

.52) compared to those whoscored below 55% (*M* = 3.65, *SD* = .58). The difference was significant, since *f(*2, 309) = 7.38, *p* = .000.

Table 5

ANOVA results when disaggregated by marks obtained

	Below 55%		*Above 55%*		*f(2, 309)*	*Sig*
	M	*SD*	*M*	*SD*		
Learning	3.65	0.58	3.88	0.52	7.38	0.00
Enthusiasm	3.73	0.67	3.98	0.66	5.23	0.01
Organization	3.67	0.66	3.85	0.64	2.94	0.05
Group interaction	3.81	0.80	4.05	0.71	3.87	0.02
Individual rapport	3.89	0.61	4.14	0.64	5.92	0.00
Breadth	3.72	0.59	3.86	0.59	2.10	0.12
Exams	3.72	0.52	3.95	0.58	7.49	0.00
Assignments	3.80	0.60	4.06	0.67	6.31	0.00
Teacher	3.49	0.64	3.83	0.53	13.57	0.00

Similarly, students who scored greater than 55% in the last examinations perceived that their teachers were more enthusiastic about teaching them (*M* = 3.98, *SD* = .66) compared to those students who scored less than 55% (*M* = 3.73, *SD* = .67), and that the difference was significant at *f(2, 308)* = 5.23, *p* = .01. In a similar manner, students who scored greater than 55% perceived that their teachers organized their lessons well (*M= 3.85, SD = .64) compared to students who scored less than 55% (M = 3.67, SD = .66).* Thedifference was significant at *f(2, 309)* = 2.94, *p = .05.* Similar differences were obtainedfor group interaction, individual rapport, examinations, assignments, and perceptions of their teachers.

Discussions

Classroom teaching has the greatest influence on student learning [38, 39]. The results of this research indicated that students had favourable view of the quality of education that they received in schools. Participants reported that despite the conceptual challenges, they were able to learn the concepts well, indicating a developmentally appropriate curricular objective. Similarly, participants also reported that their teachers were enthusiastic about teaching and

organized the presentation of the contents well. However, research has also established that Eastern countries are collectivists in nature and it is against the accepted norms to challenge the authority and the prestige of the teachers [40, 41, 42, 43]. Therefore, to what extent do the reported scores represent the actual perceptions of the students is questionable. The results also suggested that the teachers facilitated group interactions in their classroom and had a good rapport with the students. Similarly, participants reported that teachers covered the breadth of the topics taught, had a positive view of examinations and assignments, along with favourable views of their teachers. These findings contradict the findings of previous research in Bhutan. For example, previous research in Bhutan had determined that teaching-learning process was teacher centered [44], examinations was detrimental to authentic student learning [45], and that student learning was two grades below the expected level of competencies required [46]. Therefore, conforming to previous research findings, the validity of assessing education quality based on student assessment definitely seems questionable or lacking validity [47, 48, 49, 50].

Results also suggest that there are no gender differences in the quality of education that the students receive except for group interaction and individual rapport. Girls had a more positive perspective on both group interaction as well as individual rapport with the teachers. However, disaggregating students' performance based on the marks obtained in the previous examination, it was also evident that even those students who scored less than 55% in the last examinations had a positive view of the various dimensions of education quality. Therefore, these findings also question the validity of the scores obtained.

Limitations and future research

This research used cross-sectional survey data to determine students' perspectives of education quality. Surveys facilitate generalizability of the findings but do not provide in-depth information about the problems being investigated. Thus, the primary limitation of this research stems from the lack of in-depth understanding of the problem. Future research should attempt to investigate the research problem with qualitative or mixed-methods approach to glean more information besides the structured interviews. Culturally, it is unacceptable for a learner to provide negative views about your teachers in the Bhutanese context. Another area for investigation for future research would be to determine how culture influence student evaluation of education quality that they receive.

Conclusion

This research found that students had a favourable view of the quality of education that they receive. Participants reported that they were learning despite the concepts being challenging, and that the teacher was enthusiastic about teaching and were able to organize the delivery of the lessons well. Similarly, participants reported a positive perspective on group interaction, individual rapport, breadth of the topics covered, examinations, assessment, and their teachers. However, cultural influences in students scores are also evident, rendering the validity of the scores questionable. It is recommended that further investigations be conducted to determine participant's authentic assessment of education quality to inform policies and practices.

References

[1] Abrami, P. C. (1989) How should we use student ratings to evaluate teaching? *Research in Higher Education, 30,* (2), pp. 221–227.

[2] Hellman, C. M. (1998). Faculty evaluation by students: A comparison between full-time and adjunct faculty. *Journal of Applied Research in the Community College, 6(*1), 45-50.

[3] Marsh, H.W., &Roche, L. (1993). The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal, 30*(1), 217–251

[4] Chen, Y., &Hoshower, L.B. (2003). Student Evaluation of Teaching Effectiveness: An assessment of student perception and motivation. *Assessment & Evaluation in Higher Education, 28* (1), 71-88.

[5] Ministry of Education, (2021). *Bhutan professional standards for teachers.* Ministry of Education, Royal Government of Bhutan.

[6] Titus, J. J. (2008). Student ratings in a consumerist academy: leveraging pedagogical control and authority. *Sociological Perspectives, 51*(2), 397–422

[7] Fairchild, E., &Crage, S.M. (2014). Beyond the debates: measuring and specifying student consumerism. *Sociological Spectrum, 34,* 403 - 420.

[8] Molesworth, M., Nixon, E.,& Scullion, R. (2009). Having, being and higher education: the marketisation of the university and the transformation of the student into consumer, *Teaching in Higher Education, 14*(3), 277-287, DOI: 10.1080/13562510902898841

[9] Baxter, E. P. (1991) The TEVAL experience, 1983-88: The impact of a student evaluation of teaching scheme on university teachers, *Studies in Higher Education, 16,* 151-178.

[10] Arthur, L. (2009). From performativity to professionalism: lecturers' responses to student feedback, *Teaching in Higher Education, 14*(4), 441-454, DOI: 10.1080/13562510903050228

[11] Lang, W. B. & Kersting, M. (2007). Regular feedback from student ratings of instruction: Do college teachers improve their ratings in the long run? *Instructional Science, 35*(3), 187-205.

[12] Blair, K., & Noel, K.V. (2014). Improving higher education practice through student evaluation systems: Is the student voice being heard? *Assessment & Evaluation in Higher Education, 39,* 879-894. https://doi.org/10.1080/02602938.2013.875984

[13] Richardson, V. (2005). The diverse learning needs of students. In D. M. Billings, & J. A. Halstead (Eds.), *Teaching in nursing* (2nd ed.). Elsevier.

[14] Kember, D., Leung, D., & Kwan, K-P (2002). Does the use of student feedback questionnaires improve the overall quality of teaching? *Assessment & Evaluation in Higher Education, 27,* 411-425. 10.1080/0260293022000009294.

[15] Schmelkin, L. P., Spencer, K. J., & Gellman, E. S. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education, 38*(5), 575-592.

[16] Wong, W. Y. & Moni, K. (2014). Teachers' perceptions of and responses to student evaluation of teaching: purposes and uses in clinical education, *Assessment & Evaluation in Higher Education, 39(*4), 397-411, DOI: 10.1080/02602938.2013.844222

[17] Yao, Y., & Grady, M. (2006). How do faculty make use of student evaluation feedback? A multiple case study, *Journal of personal Evaluation in Education, 18 (2),* 107-126. DOI:10.1007/s11092-006-9000-9.

[18] Beran, T. N., &Rokosh, J. L. (2009). Instructors' perspectives on the utility of student ratings of instruction. *Instructional Science, 37*(2), 171–184. https://doi.org/10.1007/s11251-007-9045-2

[19] Golding, C.,& Adam, L. (2016) Evaluate to improve: useful approaches to student evaluation, *Assessment & Evaluation in Higher Education, 41*(1), 1-14, DOI: 10.1080/02602938.2014.976810

[20] Shakurnia, A. H., Mozafari, A. R., & Karami, M. A. (2012). Comparison of teachers' attitude with high and low scores evaluation on the importance and use of student assessments. *Educational Research in Medical Sciences, 1*(1), 38–46.

[21] Simpson, P., &Siguaw, J. (2000). Student evaluations of teaching: an exploratory study of the faculty response, *Journal of Marketing Education 22*(3), 199-213. DOI:10.1177/0273475300223004

[22] Stein, S. J., Goodchild, A., Moskal, A., Terry, S., & McDonald, J. (2021). Student perceptions of student evaluations: enabling student voice and meaningful engagement, *Assessment & Evaluation in Higher Education, 46(*6), 837-851, DOI: 10.1080/02602938.2020.1824266

[23] Zabaleta, F. (2007). The use and misuse of student evaluation of teaching. *Teaching in Higher Education, 12*, 55–76. doi:10.1080/13562510601102131

[24] Seldin, P. (1993). The use and abuse of student ratings of professors. *Chronicle of Higher Education, 39,* A40.

[25] Penny, A. R. (2003). Changing the agenda for research into students' views about university teaching: Four shortcomings of SRT research. *Teaching in Higher Education, 8,* 399–411. doi:10.1080/13562510309396

[26] Balam, E.M., & Shannon, D.M. (2010). Student ratings of college teaching: a comparison of faculty and their students. *Assessment & Evaluation in Higher Education, 35*, 209 - 221.

[27]. Penny, A. R., & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research, 74*(2), 215–253. https://doi.org/10.3102/00346543074002215

[28] Douglas, J., & Douglas, A. (2006). Measuring student satisfaction at a UK university. *Quality Assurance in Education 14*(3), 251-267. DOI:10.1108/09684880610678568

[29] Ory, J. C. (2001). Faculty thoughts and concerns about student ratings. *New Directions for Teaching and Learning, 87*, 3–15.

[30] Hutchinson, S. R. (2003). Survey research. In S. deMarrias, & S. D. Lapan (Eds), *Foundations for research: Methods of inquiry in education and social sciences* (pp. 283-301). Lawrence Erlbaum Associates, Inc.

[31] Creswell, J.W. (2013).*Research design: qualitative, quantitative, and mixed methods approaches*. (4th ed), Sage.

[32] Creswell, J. W. & Creswell, J. D. (2017). *Research design: qualitative, quantitative, and mixed-methods approaches* (4th ed). Sage Publishers.

[32] Mertens, D. M. (2005). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (2nd ed). Sage Publications.

[33] Mertens, D. M. (2015). *Research and evaluation in education and Psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (4th ed). Sage Publications.

[34] Fraenkel, J. R., & Wallen, N. E. (1990). *How to design and evaluate research in education*. McGraw-Hill.

[35] Fowler, F. J. (2014). *Survey research methods* (5th ed). Sage Publications.

[36] Polit, D. F. & Beck, C. T. (2010). Generalizations in qualitative and quantitative research: Myths and strategies, *International Journal of Nursing Studies, 47*, 1451-1458. doi:10.1016/j.ijnurstu.2010.06.004

[37] Nunnally, J. C. & Bernstein, I. H. (1994). The assessment of reliability. *Psychometric Theory, 3*, 248-292.

[38] Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership, *School Leadership and Management, 28*(1), 27-42. DOI:10.1080/13632430701800060

[39] Leithwood, K., Harris, A., & Hopkins, D. (2019). Seven strong claims about successful school leadership revisited, *School Leadership and Management, 40*(4), 1-18. DOI:10.1080/13632434.2019.1596077

[40] Hofstede, G. (2011). Dimensionalizing cultures: the Hofstede model in context. *Online Readings in Psychology and Culture, 2*(1). https://doi.org/10.9707/2307-0919.1014

[41] Löckenhoff, C. E., Lee, D. S., Buckner, K. M. L., Moreira, R. O., Martinez, S. J., & Sun, M. Q. (2014). Cross-cultural differences in attitudes about aging: Moving beyond the east-west dichotomy. In Cheng, ST., Chi, I., Fung, H., Li, L., Woo, J. (eds) *Successful Aging*, (pp. 321–337). Springer. doi:10.1007/978-94-017-9331-5_19

[42] O'Brien, E. L., Hess, T. M., Kornadt, A. E., Rothermund, Klaus, Fung, H., & Voss, P. (2017). Context influences on the subjective experience of aging: The impact of culture and domains of functioning. *The Gerontologist, 57*(suppl_2), S127–S137. doi:10.1093/geront/gnx015

[43] Reich, A. J., Clauch, K. D., Verdeja, M. A., Dungan, M. T., Anderson, S., Clayton, C. K., Goates, M. C. & Thacker, E. L. (2020). What does "successful aging" mean to you?- Systematic review and cross-cultural comparison of lay perspectives of older adults in 12 countries, 2010-2020. *Journal of Cross-Cultural Gerontology, 35*:455–478. https://doi.org/10.1007/s10823-020-09416-6

[44] Royal Education Council (2009). *The quality of school education in Bhutan: Reality and opportunities*. Royal Education Council.

[45] Johnson, D., Childs, A., Ramachandaran, K., & Tenzin, W. (2008). *A needs assessment of science education in Bhutan.* Curriculum and Professional Support Division, Ministry of Education.

[46] Education Sector Review Commission (2008). *Education without compromise*. The National Council of Bhutan.

[47] Feistauer, D. & Richter, T. (2017) How reliable are students' evaluations of teaching quality? A variance components approach, *Assessment & Evaluation in Higher Education, 42(*8), 1263-1279, DOI: 10.1080/02602938.2016.1261083

[48] Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M. T., Filer, J. D., Wiedmaier, C. D., & Moore, C. W. (2007). Students' perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal, 44,* 113–160. doi:10.3102/0002831206298169

[49] Spooren, P., Brockx, B., &Mortelmans, D. (2013). On the validity of student evaluation of teaching: the state of the art. *Review of Educational Research, 83*(4), 598–642. https://doi.org/10.3102/0034654313496870

[50] Spooren, P., & Christiaens, W. (2017). I liked your course because I believe in (the power of) student evaluations of teaching (SET): students' perceptions of a teaching evaluation process and their relationship with SET scores. *Studies in Educational Evaluation, 54,* 43–49. https://doi.org/10.1016/j.stueduc.2016.12.003

[51] Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press