Innovation in online Evaluation for authentic assessment

Dr. Lochan Jolly[1], Mr. Arpit Vyas[2], Mr. Swapnil Raut[3] and Dr. B.K. Mishra[4] Dean SSW & Professor[1], Assistant Professor[2,3], Principal & Professor[4] Thakur College of Engineering & Technology, Kandivali, Mumbai

***Abstract*—In the present pandemic condition, challenge as a teacher, is not only to learn new tools for online teaching but also conducting online assessment and to assure that the assessment curves are not flat but shall be a natural response which has bell shaped distribution. We all know the impact of flat evaluation is that the industry will lose trust on the quality of graduate and their knowledge based on marks/grades. This will put the quality of technical education under doubt when high scorers perform poorly in the recruitment process (including job interviews) where students are not even able to pass the basic aptitude clearance test.**

Online education may not lack the quality and integrity of face-to-face instruction, but it is easy to cheat in online assessment. So, there is a need to develop some innovative and best practices which will help the sincere students to have academic edge over the insincere students and the students' marks or grades shall be the replica of their intellectual ability. It will enable the deserving students to get their due credits and not bring all at the same level. Current problem of flat evaluation which has become more significant in the current pandemic can be overcome. This is mainly because of online examinations in the home environment. This paper suggests and tests some techniques to remove flatness in evaluation and make it more authentic such as, use of uniqueness and uncertainty in question paper and assessment, conducting multiple activities as part of evaluation, giving weightage to various activities as per the efforts, making assessment a multilevel activity etc. The results show the evaluation curves which were flat when no innovative assessment technique was used for assessment became bell shaped when innovative assessment technique was used. Moreover, it was found to be in alignment with testing based on classroom testing under a proctored environment.

***Keywords*—Innovation, Evaluation, Assessment, Normal distribution**

INTRODUCTION

The government guidelines to prevent the spread of coronavirus (COVID-19) has impacted all the educational institutes as students and faculty cannot come to campus and all the classes are being conducted through online mode. All universities, institutions and other agencies around the world are trying to find tools and technology for conducting the authentic assessment in the home environment so that social distancing measures can be followed. These exams are commonly used to make high-stakes decisions. Some of these decisions related to institutions for students include:

- Promotion of students to the next education level
- Certifying Student's eligibility for professional career and the prediction for their professional career.
- Certifying student completed course as per course duration with subject backlog or no backlog.
- Certifying students' knowledge, skills and attitude as they enter the professional career
- Quantitative results are required to be converted to qualitative results defined in terms of graduate attributes.

Thus, the results of such exams directly affect students' career, livelihood and futures opportunities. When the examination is properly designed and administered, it can enhance equity by opening up access to educational and placement opportunities for students from diverse backgrounds with inclusiveness. Likewise, high-stakes examinations can also increase transparency, accountability, and integrity in the student certification or selection processes.

Till the country was hit by pandemic, high-stakes exams tended to be administered in person, in paper or online format, with students and proctors physically present together in an institute or a testing center. In the current context, institutions, universities, governments and testing organizations have to taking into account health and safety considerations while setting exam administration guidelines and also evaluate the extent to which it is feasible. Moreover, mobility of the individuals is also restricted due to stringent implementation of COVID SOP including lockdown where the spread is uncontrolled.

Under this condition to continue with the examination and evaluation & assessment in traditional way is not possible and setting altogether a new system for the same is a challenging task. Now a days institutions are required to invest a large amount to create the digital platform for solutions so that digital

delivery of learning content and on administering highstakes examinations online can be done. This will allow students to learn and take exams from their homes without risking exposure to novel coronavirus. The challenge of online examination is assuring the authenticity of evaluation and assessment.

Some of the exam solutions with digital platform used are as follows:

1. G Suite Google form (Free)
2. Microsoft teams (Free)
3. Yaksh (Free)
4. Exam.net
5. Parikshit

Some problems when using free software is students can easily collaborate and the evaluation is flat curve which may raise questions about the authenticity of the evaluation system. This is mainly because of sharing of answers because of the following reason.

- Physical monitoring is not possible as the examination is conducted in the home environment where students are in possession of the class notes, textbook, reference book (including in PDF format), faculty ppts, video link etc. Online contents are also abundant as testing is mainly for low order skills.
- Some of the faculty members started using the platform in the last few months may not be able to exploit the full features of software and therefore maturity for usage will take some time. Same situation is applicable for service providers also as online examination was not very popular in the past for such high-stake examinations.
- Usage of google platform for searching the answers as the students in the university system are tested based on frequently asked questions and therefore lack uncertainty and the uniqueness of the question paper. Therefore, the answers are readily available on google.
- Usage of anti-proctoring software tools that nullify the technology-based proctoring and the remote human based proctoring through video link may not be very efficient and effective.
- Some students are found to take advantage of technical glitches.

As a result, those who are working hard and with sincerity, score low as they choose not to cheat. Hence first time they must have opted to be honest but next time they will also take the shortcut to score high using cheating and collaboratively solving the papers. If all students get 100% score it will create doubt about the credibility of the institute and evaluation system among the recruiters. It may be dangerous for the future of students of batches to come. This paper suggests and tests some of the ways to enhance the online evaluation system to make it authentic and performance of the students can be fitted in to the normal distribution which of the need of the hour

I. FINDINGS

From the recent past several months, the global spread of the current pandemic due to the COVID 19 virus has forced the people to remain in the house and also to work from home as far as possible. In the process, to ensure the educational institution has adopted an online Teaching-Learning Process (TLP), Examination and Evaluation in home environment. Technology has improved a lot in such segments. Lots of online tools are also made available and the products are improving on a continuous basis due to regular innovation in the field. However, technology is required to be used with honesty, transparency, accountability, integrity and ethics which is found missing. As a result, malpractice in the examination viz using google platform, sharing of answers with multiple devices, usage of anti-proctoring tools etc. during the examination has increased. Flat high marks from evaluation have become very common in the system. Moreover, post COVID, online education in the blended mode is going to be the new normal. Looking into need, the author has taken the project to understand the problem and come out with the solution so that the traditional examination system can be supplemented with a strong Innovative Examination System. It will enable the system not only to correct the system where marks or grades obtained by the students are in proportion with their actual knowledge acquired but also upgrade the traditional system for future need of education. Evaluation and assessment (E&A) load for faculty members are going to be at par with TL load. It will lead to true education where graduates will be professionally, ethically and socially competent. To understand the E&A best and innovative practices around the world and their acceptance to make the online evaluation authentic a rigorous study and survey was carried. The study included literature survey, educational survey and analytical study. The literature survey helped to understand the tools and techniques used around the world for online examination, evaluation and assessment and educational survey from students and faculty helped to understand the acceptance of the best practices of online examination, evaluation and assessment. Analytical study of survey's helped to device a methodology which had higher acceptability by students and faculty and could give authentic results for E&A of student's knowledge.

A. Literature Survey

Traditional exams represent the most appropriate assessment tool, but there are ways to enrich authenticity. Assessment techniques such as case studies, study papers, scenario-based projects, word problems, video assignments, open book test, short-answer items, simply asking students to record a brief spoken-word explanation of their answers and return it after the exam etc. can often represent a superior way of measuring students' learning, engaging and empowering them to demonstrate knowledge rather than demanding that they prove their worth via high-stakes exams(*How to Discourage Student Cheating on Online Exams (Opinion)*, n.d.), (*Designing Effective Online Assessment | Teaching & Learning - UCL – University College London*, n.d.) These innovative techniques are effective and can be used instead of traditional evaluation examination during this pandemic and are expected to give authentic results for online evaluation and assessment (Chavan et al., 2018) [3-5]. The developments in assessment using the online tools would only be authentic if they take account of the way the online tools functions outside of higher education, rather than seeing it as an educational technology divorced from its own authenticity.[7]. The objective of the whole exercise of identifying tools for assessment is students should motivated students to learn. Brophy et al stresses that there is a need to identify and outlines numerous pointers for creating essential learning preconditions and for motivating, including maintaining high expectations, supplying extrinsic incentives, and capitalizing on students' intrinsic motivation and also provides strategies for stimulating students to learn course content.[8]. B. Arend et al. highlights, effective assessments practice by using multiple and alternative assessment methods, dispersing grades over time, and providing timely and frequent feedback to students. For enhancing students learning he also focuses on student's report on relatively more complex learning strategies, such as elaboration and critical thinking over rehearsal. He even warns that online instructors need to ensure that assessments are used strategically, and that feedback is productive and able to be acted upon by student.[9]. Dawn Birch et al. shared students experience regarding online assessment and also highlights that it allowed students to achieve a range of cognitive and social learning outcomes, as well as to develop some important graduate skills.[10]. Christine Greenhow et al. did analysis of his experiments with online activities and shared that technologies can be incorporated into the students' overall learning ecology to reduce educational inequities and current institutionalized approaches must shift to accommodate change [11]. The use of authentic assessments contributed to improving credibility at the program and institutional levels is highlighted by Lilia Juele et al. He also stresses on the need for institutions to support faculty in developing assessments for online courses by providing the needed incentives, training and professional development that will help instructors gain fluency in the design and development of authentic assessment that will equip online students with critical thinking skills necessary to meet the challenges they will face after graduation [12].

Ence Surahman et al. shares peer assessment approach. According to him Peer-authentic collaborative assessments provide opportunities for students to assess their learning progress. Student comes to know that the peer collaborative authentic assessment is deemed necessary and relevant to the evaluation characteristics of 21st century learning that is oriented towards improving higher order thinking skills as an effort to build student capabilities [13]. Dr. Wendy Barber at el. highlights shift in the nature of knowledge and its mismatch with the nature of assessment. He shares the importance Problem based learning to shape students' knowledge, and acquire the key attitudes necessary for success in a digital world. He rightly points outs the need to authentically assess students by checking their ability to demonstrate their knowledge in a variety of artistic and creative ways that best fit their digital skills and knowledge, and should develop the confidence and competence to participate in meaningful online communities [14]. Richard Osborne at el. shares a new model for ''authentic'' assessment design, which integrates off-the-shelf technologies using an affordances approach. That model suggests that it is effective in supporting the design of an ''authentic'' assessment and that the targeted affordances approach can support the alignment of technologies with a pedagogic design [15]. According to Todd J. B. Blayone at el. the Fully Online Learning Community model is a community-oriented, digitallearning model guided by a strong democratic orientation, and a commitment to socially and cognitively rich, collaborative learning. He appreciates that the online learning platform is not only solution to present condition problems but also provides opportunities [16]. Mischelle Taylor Stone at el. highlights the benefits for traditional and its ability to enhance the quality of students learning [17]. Miller at el. compares online learning and MOOCs platform [18]. Olfos, R. at el. focuses on the improvement in evaluation instruments, system's design, refined rubrics, and rigorous quality criteria to make assessment authentic [19]. Maria Hanifah at el. dicusses the role of the government to train faculty for designing authentic assessment so that the teachers will be able to understand the essence within the implementation.

[20].

The literature survey gave an insight about needs of online evaluation system and also gave an idea about methods that can be used to evaluate the learning of students in online evaluation and assessment.

Inputs from literature were taken for suggesting the new methodology suitable for the present education system in higher and technical institutes which will help to remove flatness in online evaluation and assessment.

B. Educational Survey

Based on literature survey we designed an educational survey [6] to understand the acceptability of identified changes in online evaluation model from students and faculty. 489 people including students and faculty responded. The survey helped to gather feedback and suggestions from primary stakeholders about existing present form and suggested future form of online examination, evaluation and assessment based on literature. The study showed that students and faculty approve that the identified techniques for online evaluation and assessment will give authentic results.

C. Analytical study

The literature survey and educational survey provided data base to understand the cause of failure of online examination, evaluation and assessment. The analytical study was done using two tools.

1. *Root Cause Analysis using 5 why model*. - To analyze the real cause for flat evaluation in online assessment we did Why-Why analysis with the senior leadership quality circle team consisting of HODs, Dy. HODs and coordinators

TABLE I

SUMMARY OF 5 WHY ANALYSIS

Question	Answer
1. Why do we have a flat evaluation response?	Because they can easily collaborate and cheat in online mode and hence all get full marks in MCQ and also all attempt same questions in subjective questions.
2. Why do students cheat?	Because they have exam pressure and stress of failing and attaining low grades.
3. Why do they have pressure and stress?	Because they do not feel connected, supported and encouraged being part of the student by being honest.
4. Why students do not feel connected, supported and encouraged in online system?	Because evaluation techniques do not use tools which invites them to demonstrate learning in ways that are most authentic to them.
5. Why there is a need to make E&A authentic?	Because the present system is giving Flat evaluation curve.

2. *Fishbone analysis* - Further Fish bone analysis (fig 1) was done with the senior leadership quality circle team consisting of HODs, Dy. HODs and coordinators.

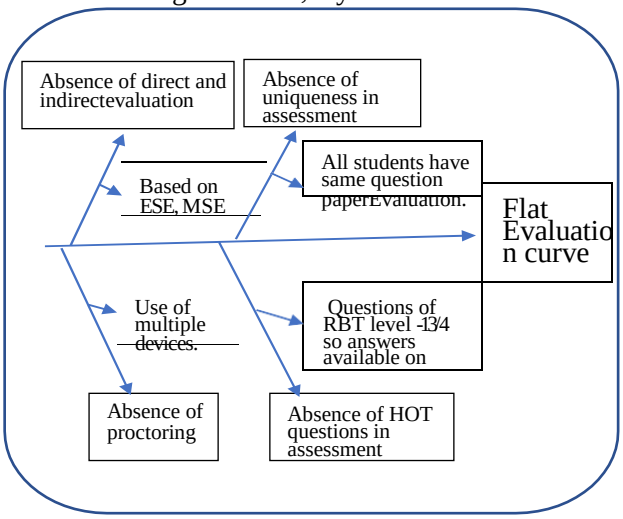


Figure-1 – Fish bone analysis

3. It was found that the factors for the flat evaluation curve along with their weightage are (I) Absence of uniqueness in assessment (25%): All students have the same question paper to answer and hence they can easily collaborate and get answers for all the questions (II) Absence of indirect evaluation (25%): Traditional Evaluation Techniques are used for evaluation under online mode. The final grades are as per end semester evaluation and in semester evaluation (mid semester evaluation). So, there is absence of direct and indirect evaluation to evaluate students actual learning through other activities like their involvement during sessions, their responses during sessions, their ability to give solutions for various problems related to topic etc. (III).

Absence of proctoring (25%): It is easy to collaborate under online mode. As most of the students have multiple devices so even if we want to proctor them during exam, they use another device to collaborate and get the answers from the peers or copy from google. (IV) 4. Absence of Higher order thinking questions in assessment (25%): Questions of assessment are of RBT

level 1-3 or 4 so students can answer easily find by doing google search.

D. Suggested Methodology

Based on the rigorous study, we propose Multiple level evaluation to enhance the authenticity of online evaluation. The proposed methodology does evaluation in following five levels.

TABLE II

SUGGESTED LEVELS OF EVALUATION

Level	Type of question	Weightage	RBT level	Platform
1	MCQ	30%-40 %	1	Google forms
2	Single word/ short answers	40 %-30%	2	GCR
3	Long questions/ Numerical	20 %-30%	2 & 3	GCR
4	Case study and practical related analysis questions	10%	4 & 5	GCR
5	Moderation and presentation	Multiplication factor	1 to 4	Zoom

Following guidelines can be used for designing multilevel

Evaluation and assessment

a) The uniqueness and uncertainty can be created by taking the MCQ of all subjects together. That is, we could have one final paper of 1.5 hrs. duration and 60 MCQ in addition to the traditional technique of evaluation.
b) Multiple sets of question paper for all students should be made using the software and question bank so that all students do not get same paper.
c) MCQ can have questions with multiple correct options.
d) Divide paper into sections (for each subject) and subsections (within subject) along with shuffling and make submission of each section/ sub section mandatory before going to next section/sub section. This will further add redundancy and hence reduce copy and collaboration.
e) Multiple choice questions related to one case study or practical can be added.
f) The uniqueness and uncertainty can be created in numerical by giving one unique parameter for each student may be roll no. so that each student gets unique answers and students can be asked to type answers instead of giving option to select from the choices.
g) For a given case study student can be asked to design numerical and also provide solution for it.
h) Students can be asked to give presentation as a part of evaluation.
i) Rubrics given in the table 3 can be used for evaluation and assessment.
j) After all the exams a Viva can be conducted to test the student's real knowledge for all subject to decide the multiplication factor for each student to calculate his actual marks. This will motivate the students to study in detail from the books. It will also help to identify the correct learning level of the students.
k) In online mode for conducting all the suggested activities we can use different platforms for different level of evaluation and assign weightage to each activity as shown in table II to get bell shaped evaluation curve.
l) MCQ can be taken on google form, Design or open book or HOT questions student can submit on google classroom, students can give presentation on any subject topic or case study on zoom classroom

TABLE III

RUBRICS FOR EVALUATION

Theoretical question	Numerical questions	% Marks
Wrong answer	Wrong answer	0%
Partially right answer with diagram,	Partially right steps	Mention percentage
All points covered with Language, clarity, articulation, and decimation.	All steps covered and diagram with everything correct and drawn to scale and labelled correctly	80%
Beyond content covered like mention of examples related to recent case studies	Innovation in solving. Like new method or short method	20%

E. Results and Discussion

Based on proposed methodology for evaluation a formative Assessment (FA) for the subject of Fiber optic communication (SEM VI professional elective) was designed. The FA was conducted using Nearpod tool activities and the evaluation had two parts first part was a simple MCQ (10 questions) where students were given 1 minute for answering each question and the leader board scores were displayed after each question to motivate

students. Students had only one chance to attempt one question and after 1-minute next question was displayed.

Second part was open ended questions (5 questions), in which students with understanding of the topic could answer. It was a one-line short answer and also a timed activity with 1.5 minutes for each question Fig 2 shows that scores in traditional system of evaluation which was conducted in proctored and controlled environment of the class. It can be seen from the plot that it was giving the bell-shaped evaluation curve. The reason for this performance is that, because of proctored environment students cannot collaborate and have to write answers as per their knowledge and hence we could get variation in grades

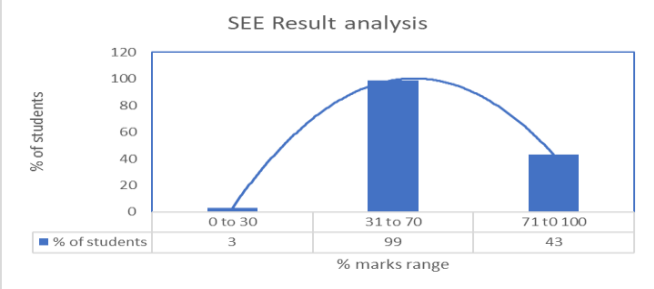


Figure -2- Result analysis of SEE for Optical communication (exam was conducted in physical form) Fig 3 shows that the evaluation curve is flat when the assessment was as per traditional evaluation method in online mode. The reason for this performance is that students can collaborate easily and hence get same grades

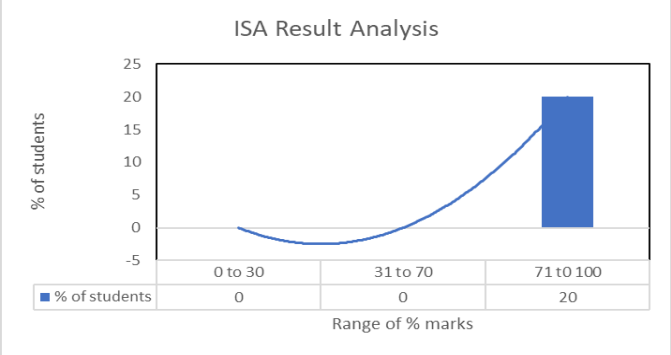


Figure -3Result without innovative techniques for online exam

Fig 4. shows there was bell shaped variation in evaluation curve when innovative techniques used for assessment. The plot shows that the final scores for the students was not flat but followed the bell-shaped curve with 15% students in bottom, 5 % students in top and remaining 75 % students in middle. This was because students had to answer the questions as per individual's knowledge as there was time limitation and less time to collaborate, so they got different grades.

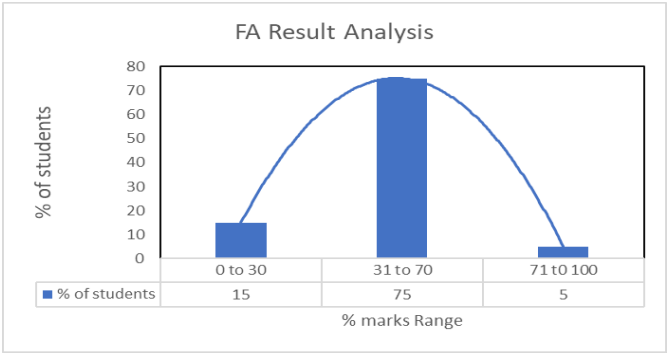


Figure-4- Result with innovative techniques for online exam

F. Conclusions

The challenges of online exam are not trivial. This paper does a detailed analysis to identify the root cause for flat curve for evaluation and assessment in online E&A. The root cause identified for flat response in the online E&A is lack of uniqueness & uncertainty in evaluation paper and the ease to collaborate during assessment for the students in online mode. We propose an online system for assessing students' knowledge. In this paper few methodologies for creating uncertainty and uniqueness in online evaluation were suggested and also tested. The test results show that the suggested methodologies can help to make online E&A authentic as we are getting a bell shape curve for E&A in online mode as we used to get for E&A when it was conducted in the controlled and proctored environment in offline mode. This is because E&A with new ideas reduces the chances to collaborate as it will create uncertainty and uniqueness.

For adopting this practice in system, it needs to be calibrated with Teaching Learning during doubt solving and practice tests so that everything can streamlined and do not come as a surprise for the students.

Thus, this system can help us to evaluate the actual learning for the students as it will reduce malpractice and will lead to bell shaped evaluation curve.

G. References

1. https://www.insidehighered.com/advice/2020/04/29/how-discourage-student-cheating-online-exams-opinion
2. https://www.ucl.ac.uk/teaching-learning/assessmentresources/designing-effective-online-assessment
3. Deepti Chavan, Bhushan Nemade, Bhavika Nemade," Methodologies for Enhancing the Quality of Exam

Question Paper Setting and its Evaluation", National Conference on Evaluation & Assessment Reforms (NCEAR), pp 53-59,2018.

4. Lochan Jolly, Rutvi Thakkar," Impact of Evaluation System on Student's Learning Pattern", National Conference on Evaluation & Assessment Reforms (NCEAR), pp 18-23,2018.
5. Prachi Janrao, Ruta Pathak," Evaluation & Assessment Tools in Engineering Education", National Conference on Evaluation & Assessment Reforms (NCEAR), pp 9-11,2018.
6. Looney, J. W., "Assessment and Innovation in Education", OECD Education Working Papers, No. 24, OECD,2009
7. Matthew Allen," Authentic Assessment and the Internet: Contributions within Knowledge Networks"

Association for the Advancement of Computing in Education (AACE), San Diego, CA,Oct 26, 2009 in Vancouver, Canada ISBN 978-1-880094-76-1. 8. Brophy, Jere, "Synthesis of Research on Strategies for Motivating Students to Learn" Educational Leadership, v45 n2 p40-48 Oct 1987.

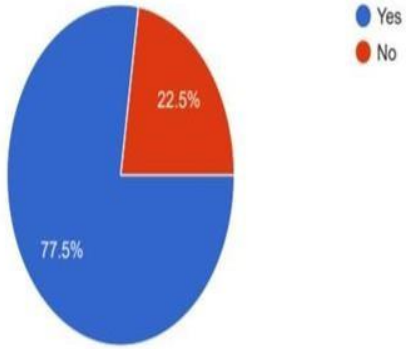
9. B. Arend, "Course assessment practices and student learning strategies in online college courses" The Online Learning Consortium's, Published in 2007.
10. Dawn Birch, Michael Volkov," Assessment of online reflections: Engaging English second language (ESL) students" Australasian Journal of Educational Technology Volume 23, Number 3, Jan 01, 2007 ISSN 0814-673X Publisher: Australasian Society for Computers in Learning in Tertiary Education.
11. Christine Greenhow, Beth Robelia," Informal Learning and Identity Formation in Online Social Networks". Learning, Media and Technology Volume 34, Number 2, June 2009 ISSN 1743-9884.
12. Lilia Juele, Rockland Community College (SUNY), United States," Authentic Assessments: A Critical thinking and Engagement Tool for Online Courses" Association for the Advancement of Computing in Education (AACE), Waynesville, NC, Jun 25, 2018.
13. Ence Surahman, Agus Wedi, Sulthoni, Yerry Soepriyanto, Punaji Setyosari, "Design of Peer Collaborative Authentic Assessment Model Based on Group Project Based Learning to Train Higher Order Thinking Skills of Students". Advances in Social Science, Education and Humanities Research, volume 285. International Conference on Education and Technology (ICET 2018).
14. Dr. Wendy Barber, Sherry King and Sylvia Buchanan," Problem Based Learning and Authentic Assessment in Digital Pedagogy: Embracing the Role of Collaborative Communities" The Electronic Journal of e-Learning Volume 13 Issue-2, 2015.
15. Richard Osborne, Elisabeth Dunne and Paul Farrand," Integrating technologies into ''authentic'' assessment design: an affordances approach". Research in Learning Technology, Vol. 21, 2013, Coaction Publishing.
16. Todd J. B. Blayone, Roland vanOostveen, Wendy Barber, Maurice DiGiuseppe and Elizabeth Childs," Democratizing digital learning: theorizing the fully online learning community model". Blayone et al. International Journal of Educational Technology in Higher Education (2017) 14:13. DOI 10.1186/s41239-017-0051-4.
17. Mischelle Taylor Stone, Suzanne PerumeanChaney," The Benefits of Online Teaching for Traditional Classroom Pedagogy: A Case Study for Improving Face-to-Face Instruction". MERLOT Journal of Online Learning and Teaching. Vol. 7, No. 3, September 2011.
18. Miller, Sandra L.," Teaching an Online Pedagogy MOOC". Journal of Online Learning & Teaching . Mar2015, Vol. 11 Issue 1, p87-102. 16p.
19. Olfos, R., & Zulantay, H. (2007),"Reliability and Validity of Authentic Assessment in a Web Based Course". Educational Technology & Society, 10 (4), 156-173.
20. Maria Hanifah, Alfred Irambona," Authentic assessment: Evaluation and its application in science learning". August 2019. Psychology Evaluation and

ANNEXURE A

Following are the responses of new important questions of educational survey:

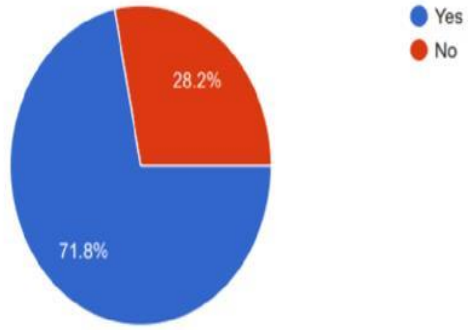
1. Do you think can we include MCQ, one-word, explanatory answers, design or case study questions and moderation for multilevel evaluation to move flatness in evaluation?

489 responses



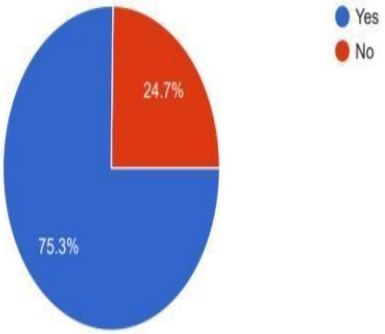
2. Do you think that this new patterns if adopted, a practice is needed to calibrate teaching learning so that everything can be streamlined and do not come as surprise for the students?

489 responses



3. Do you think this system will help us to evaluate the actual learning for the students and will lead to be shaped curve in evaluation?

489 responses



ANNEXURE B

A feedback about the innovative evaluation and assessment technique used for formative assessment.

Student	Feedback
1.	It is quite tricky to face question this way.
2.	Questions were framed better which made the test interesting.
3.	Time is limited, interesting, new experience
4.	It was kind of new, it's little bit of confusing. Time limitation was there.
5.	Something different but a little difficult to get compatible