Traditional Vs Non-Traditional Assessment Activities as Learning Indicators of Student Learning: Teachers' Perceptions

Maha Al-Anqoudi*, Mireilla Bikanga Ada*, Stephen McQuistin†, Nikos Ntarmos*, Jack Parkinson*,
and Yashar Moshfeghi‡

*University of Glasgow, UK, †University of St Andrews, UK, ‡University of Strathclyde, UK

maha.alanqoudi@gmail.com, Mireilla.BikangaAda@glasgow.ac.uk, sm@smcquistin.uk,
Nikos.Ntarmos@glasgow.ac.uk, yashar.moshfeghi@strath.ac.uk

Abstract—In online settings, some teachers express reservations about relying only on traditional assessments (e.g., tests, assignments, exams, etc.) as trustworthy instruments to evaluate students' understanding of the content accurately. A previous qualitative study revealed that the richness of online environments has allowed teachers to use traditional assessments (anything contributing to the final grade) and non-traditional assessment-based activities (not factored into the final grade but useful in gauging student knowledge) to assess their students' learning status. This study aims to compare the perceived accuracy of both types of assessment activities as indicators of student learning. A total of 124 participants engaged in online teaching completed a self-report instrument. The results revealed a significant difference in teachers' perceptions of the accuracy of traditional assessment activities (M = 3.16; SD = .442) compared to non-traditional assessment activities (M = 3.05, SD = .521), t (122) = -2.64, p = .009 with small effect size (eta = .02). No significant gender differences were observed in the perceptions of the accuracy of either assessment activities type. The most commonly employed traditional assessment activities were "final exams" (85.5%) and "individual assignments" (83.9%). In comparison, the most common non-traditional assessment methods to evaluate students' knowledge were "questions on previously taught content" (79.8%) and "asking students questions about current content during the lecture" (79%). A one-way analysis of variance revealed no significant differences in perceptions of the accuracy of traditional and non-traditional assessment activities among teachers with varying years of experience (up to 10 years, 11-15 years, and 16+ years). The findings suggest that certain non-traditional assessment activities can also be as accurate as traditional learning activities. Moreover, non-assessment-related activities are perceived to be effective learning indicators. This study has implications for academic institutions and educators interested in supplementing traditional approaches to assessing student learning with non-traditional methods.

Index Terms—Online learning, learning activities, assessments, accuracy of assessments

I. INTRODUCTION

Online teaching environments and the tools that are used have produced a new learning trend, leading to more options than in offline teaching settings. Despite the increased data that online teaching environments can collect, student learning continues to be gauged through formal assessment practices, including quizzes, assignments and exams [1]. There are many learning activities used in online teaching environments to improve the learning experience, including discussion forums [2], game-based systems [3], digital notes [4], and videos [5]. The use of such learning activities generates additional dimensions for the learning evaluation, beyond formal assessments, to assess student performance.

This study focuses on the interactions between the student and the learning content, specifically, the interactions that signal the student's understanding of the content. For example, one of the most important factors in evaluating a learner's performance is the questions that are asked to the student in the form of exams, homework, or projects, which we refer to as "Traditional Assessments Activities (TAA)" (or summative assessment activities). However, as discussed by [6] assessments are not sufficient in describing a student's actual learning status. Therefore, it is necessary to find other ways to discover the student's learning by using data generated from all possible forms of learning in the online educational environment. Teachers can discover something about the students' learning state through a variety of learning activities used during the teaching process. We refer to such learning activities as "Non-Traditional Assessments Activities (N-TAA)" (or formative assessment activities). The N-TAA are learning activities that help teachers gauge the students' learning status, some may be graded, but these grades do not count towards the course's final grade.

This paper investigates teachers' perception of online teaching in evaluating learners' learning status, looking at activities that are essential as a means for gauging students' understanding from the perspective of teachers. More importantly, this study highlights those learning activities that have not been considered in learning analytics approaches that are limited to assessments, such as exams, assignments, or projects.

II. BACKGROUND

A. Distance learning/education

To continue educating students worldwide without abandoning the value of education until the disruption from COVID-19 came to an end, the transition to online teaching and learning was necessary. As a result, a new trend started –

for many institutions – of teaching and learning in online-based environments, resulting in a different experience for both teachers and students. The educational system and educators have embraced "Education in Emergency" through a variety of online platforms, and as a result, they are now required to employ a system for which they are unprepared [7]. However, the pandemic provided a rare chance to create the foundation for evaluating digital learning in the context of complete online teaching and learning experiences. This opportunity allows us to explore the differences between traditional learning – in classrooms – and distance learning (DL). DL is defined by [8] as "the electronic means used to keep students in touch with teachers, provide access to communication between students and bridge the gap and distribute educational material through distance learning programs". DL also pertains to the learning process as well as the electronic means of that process. Building on this, [9] evaluated distance learning by looking into the "learning", "evaluation models", and "students and teachers" as the keys for distance learning evaluation. They suggested the importance of using an evaluation method compatible with the e-learning applications used.

B. Assessments for learning

In education [10], assessments are used as a factor to indicate the status of gaining knowledge of a specific concept. There are different types of activities teachers use [11] to find out whether the students understand the content or not. The activities used can be classified as traditional activities (e.g., tests, assignments, etc.) or non-traditional activities (e.g., questions and answers, group discussions, etc.). These activities (traditional and non-traditional) include even more aspects in the online learning context which requires different strategies for measuring the academic achievement of students [12]. A study [13] showed how assessments affect learning needs more broadly than feedback intervention alone by including other constructs, such as the learner's reactions to the feedback and the learning environment in which the feedback occurs. Moreover, [14] highlighted the need to properly integrate online evaluation throughout the course content to foster students' motivation. They argued that failing that integration would affect how the learning is evaluated in that course.

C. Learning process evaluation

Three types of interactions help evaluate online learning quantitively: student-to-student, student-to-teacher, and student-to-content [15]. A study in online learning [16] highlighted evidence-based approaches, including self-regulation, multimedia use in learning, motivational principles, and collaborative learning principles. Such evidence-based approaches help standardise and measure the quality of teaching and learning practices [17] in an online environment. In an online context, assessment methods, such as quizzes, are advantageous for teachers when analysing the answers [18]. However, the evaluation of learning can be influenced by other factors, such as teachers' teaching experience. The studies showed that the longer the teaching experience, the more confident teachers

are in using new techniques or teaching activities [19]. In addition, the types and numbers of learning activities used by the teachers can be different based on teachers' preferences and demographics (e.g., gender) when evaluating their students' understanding [20]. Moreover, other indicators of learning, such as attendance, proved its criticality in student progression and retention [21]. It was also shown that questions teachers ask can elicit prior knowledge and verify students' developing understanding of the content [22].

III. METHODOLOGY

A. Context of the study

The target of this study was those educators that have taught higher education courses in a full online teaching mode. Due to the COVID-19 pandemic restrictions, teachers used different online-based tools to ensure the students' understanding/tracking of the content. Moreover, teachers can choose any online-based tools in addition to what is supported in the Learning Management System (LMS). They also use a variety of activities, including traditional and non-traditional, to check whether the students have understood a concept taught.

B. Research questions

This study aims to investigate the accuracy of traditional and non-traditional assessment activities as learning indicators of students learning based on teachers' perspectives. The study also investigates the features teachers believe are effective in identifying how well students perform in the course. The following questions frame the study:

A. What are educators most used traditional and non-traditional assessment activities in online learning environments?

B. How do educators perceive the accuracy of traditional assessment activities compared to non-traditional assessment activities in online learning environments?

C. Are there any significant differences in the perceptions of the accuracy of traditional and non-traditional assessment activities among teachers with varying years of teaching experience?

D. Are there any significant gender differences in teachers' perceptions of the accuracy of traditional and non-traditional assessment activities in an online setting?

E. What features do educators believe are effective in assessing how well students perform in the course?

C. Data collection

Data was collected using a self-reported questionnaire sent on social media and via emails. Participants were people who taught online in the academic year 2021–2022. This consisted of close-ended questions classified into three sections: (i) demographics of the participants (e.g., age, gender, teaching subject, etc.); (ii) teaching experience; and (iii) mechanisms used for following up on student learning status. This third section comprised Likert scales measuring teachers' perceptions of the accuracy of TAA (10 items) and N-TAA (7 items) in evaluating students' level of understanding, and the effectiveness (7 items)

of other activities in assessing student performance. The items in TAA and N-TAA were measured on a scale of 1 to 4 (1 means "very inaccurate", 2 is "somewhat inaccurate", 3 is "somewhat accurate", and 4 means "very accurate"). The items in the effectiveness scale were rated from 1 to 4 (1 means "completely ineffective", 2 means "somewhat ineffective", 3 "somewhat effective", and 4 means "completely effective". A 4-point Likert scale was used to force participants' choices [23]. Statistical Package for Social Science (SPSS) version 28 was used to analyse data using descriptive statistics, one-way analysis of variance (ANOVA), Chi-Square tests, Paired-Samples T-test, and a Mann-Whitney U Test.

IV. FINDINGS

Participants (*n* = 124) from 16 institutions completed the teachers' survey. 60 (48.4%) were males, and 63 (51.6%) were females. Most teachers (*n*=111, 89.5%) have been teaching online for less than 3 years, 9 (7.2%) teachers have been doing so for more than three years, and 4 (3.2%) teachers would rather not say. This study was conducted during the COVID-19 pandemic, and 96.8% of the participants were actively teaching online. Out of a total of 124 responses, there were 49 (39.5%) responses for science courses, 17 (13.7%) for engineering and social science, 1 (0.8%) for medical and life science, 16 (12.9%) for arts and literature, and 23 (18.5%) for other courses.

A. What are educators' most commonly used traditional and non-traditional assessment activities in online learning environments?

Teachers were asked which traditional assessment activities (TAA) they use to evaluate the student's learning status. Results (see Table I – usage) showed that Final Exam (FE) is the most used summative assessment activity (*n*=106, 85.5%) to evaluate student knowledge in a traditional setting. This was followed by Individual assignments (IA), Mid-term exam (ME), and Class participation (CP) with 83.9% (n=104), 73.4% (*n*=91) and 61.3% (n=76), respectively. Electronic voting system (EVS) was the list used activity (*n*=15, 12.1%).

Participants (*n*=6, 4.8%) have also identified other summative activities used to evaluate student knowledge. These include report/essay writing, practical work, online exhibitions, mathematical investigation, and announced quizzes/Moodle quizzes. From the N-TAA, the top three activities that accurately help evaluate students' learning status the most are "Asking questions on previous content (QPC)", "Asking questions on current material (QCC)", and "Posts in discussion forums or online discussion area (DFP)" with 79.8% (*n*=99), 79% (*n*=98), and 56.5% (*n*=70), respectively. The N-TAA used the least was "Asking students to prepare exam style questions on the course content (SEQ)" with only 20% (*n*=25) of responses. Participants (*n*=13) also reported using other N-TAA, including google classroom for synchronous responses, Microsoft Forums, submitting drafts of the work to check during live lectures, feedback during lab sessions, use of break-out rooms for group discussions, informal feedback polls, Q&A sessions, summary

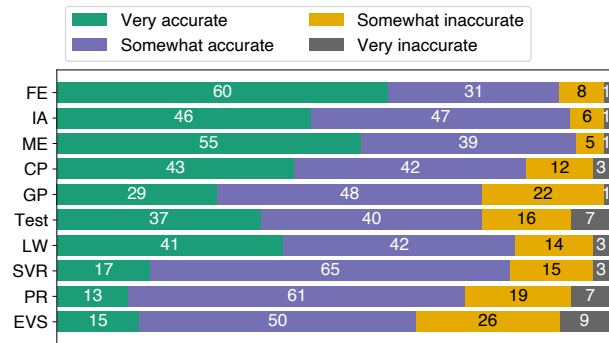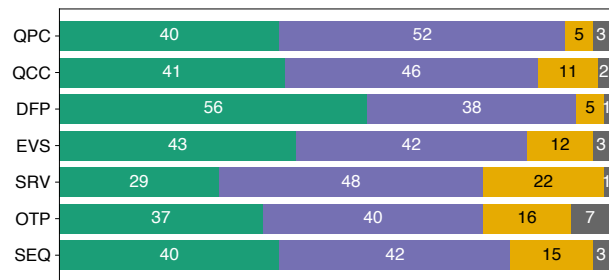Fig. 1. *Accuracy rates of TAA based on teachers' perceptions*

Fig. 2. *Accuracy rates of N-TAA based on teachers' perceptions*

report/article, screen sharing by the students to discuss their work with their teachers, and peer-to-peer help.

B. How do educators perceive the accuracy of traditional assessment activities compared to non-traditional assessment activities in online learning environments?

A paired-sample t-test revealed a statistically significant difference in teachers' perceptions of the accuracy of TAA (*M* = 3.16; *SD* = .442) compared to N-TAA (*M* = 3.05, *SD* = .521), $t(122) = -2.64$, *p* = .009 (two-tailed). The mean decrease in the accuracy scores was .110, with a 95% confidence interval ranging from .1933 to .027. The eta square statistic (.05) indicates a small effect size.

Fig. 1 shows that the most rated accurate (very accurate and somewhat accurate) TAA items are Mid exams (ME), Individual assignments (IA) and Final exams (FE) with 94%, 93%, and 91% respectively (see 1). However, discussion forum posts (DFP), Questions on previous content (QPC) and Questions on current content (QCC) received the highest rates in terms of accuracy among N-TAA items with accuracy rates of 94%, 92%, and 87%, respectively (see 2). Table I shows the usage and accuracy of the rates for both activity assessments (TAA and N-TAA). From the TAA, the final exam (FE) remains the most rated assessment activity from a usage perspective as well as the accuracy of the activities in evaluating the student's level of understanding. Likewise, QPC, QCC, and DFP are the most rated assessment activities in the N-TAA list (see Table I, N-TAA). As highlighted in (see Table I - accuracy), some N-TAA can be as accurate as TAA methods to indicate students' level of understanding in online education. Indeed,

Activity Type	Activity Name	Usage of AA		Accuracy of AA		
		N	%	N	Mean	Std. Deviation
TAA	Surprise test/quiz (Test)	50	40.3	101	**3.04**	0.927
	Individual assignments - (IA)	**104**	82.9	121	**3.39**	0.637
	Group projects - (GP)	58	46.8	106	2.92	0.885
	Peer review (e.g., using Aropa) - (PR)	16	12.9	86	2.79	0.753
	Final Exam - (FE)	**106**	85.5	120	**3.48**	0.698
	Mid-term Exam - (ME)	**91**	73.4	112	**3.46**	0.683
	Class participations - (CP)	76	61.3	114	**3.25**	0.774
	Video recording of a student demonstrating something in the course - (SVR)	30	24.2	96	2.94	0.693
	Lab work - (LW)	37	29.8	90	**3.19**	0.806
	Electronic voting system (for example,Kahoot, etc) - (EVS)	15	12.1	87	2.7	0.837
N-TAA	Asking questions on previously taught content at the start of each class(QPC)	**99**	79.8	119	**3.24**	0.701
	Asking students on current material during the lecture delivery(QCC)	**98**	79	121	**3.24**	0.753
	Posting questions/discussion topics on an online discussion area (web forum, Microsoft team, etc.) for students to address at a later time (not during the lecture)-(DFP)	**70**	56.5	106	**3.05**	0.773
	Electronic voting system (for example, Kahoot, etc) - (EVS)	36	29	91	2.89	0.823
	Asking students to record a short video explaining something from the course content (SRV)	29	23.4	92	2.99	0.832
	One-time paper where a student record their thoughts (e.g. padlet.com)-(OTP)	29	23.4	93	2.81	0.824
	Asking the students to prepare exam-style questions on the course content (e.g., using PeerWise) - (SEQ)	25	20.2	91	2.86	0.926

QPC, QCC and DFP had a mean of 3+ (somewhat accurate and very accurate), similar to Final Exam (FE), Individual Assignment (IA), and Mid-term Exam (ME), which also have a mean of 3+. Peer review (PR) and EVS were the least used assessment activities in TAA and received the lowest rates in terms of their accuracy with ($n = 86$, $M = 2.79$) and ($n = 87$, $M = 2.70$), respectively (see Table I – accuracy). Meanwhile, despite not receiving the lowest mean in terms of accuracy, "Asking the students to prepare exam-style questions on the course content (e.g., using PeerWise) - (SEQ)" was the least used non-traditional assessment activity.

C. Are there any significant differences in the perceptions of the accuracy of traditional and non-traditional assessment activities among teachers with varying years of teaching experience?

A one-way ANOVA revealed no significant differences in perceptions of the accuracy of TAA, $F(2, 120) = .99$, $p=.37$ among teachers with varying years of experience between 'up to 10 years' ($n = 44$, $M = 3.1$, $SD = .44$), '11-15 years' ($n = 32$, $M=3.2$, $SD=.42$) and '16+ years' groups ($n=48$, $M= 3.2$, $SD=.5$). There were also no significant differences in perceptions of the accuracy of N-TAA, $F(2, 121) = 2.2$, $p=.12$ among teachers with varying years of experience between 'up to 10 years' ($n=43$, $M=2.9$, $SD = .49$), '11-15 years' ($n= 32$, $M=3$, $SD=.48$) and '16+ years' groups ($n=48$, $M=3.1$, $SD=.55$). Among the items of N-TAA, a Chi-Square Test for independence indicated a significant association between teaching experience and the use of electronic voting systems (EVS), x^2 (1, $n = 107$) = 11.441, $p < .001$, phi= .363 (*medium effect size*).

D. Are there any significant gender differences in teachers' perceptions of the accuracy of traditional and non-traditional assessment activities in an online setting?

An independent-sample t-test revealed no significant gender differences in the perceptions of the accuracy of TAA for males (M=3.15, SD=.5) and females (*M*=3.17, *SD*=.4; *t*(122) =.36, *p*=.72). There were also no significant gender differences in the perceptions of the accuracy of N-TAA for males (*M*=2.97, *SD*=.6) and females (*M*=3.12, *SD*=.5; *t* (121) =1.53, *p*=.13). Results from a Chi-Square Test showed a significant difference in the proportion of females and males who use "class participation" (CP) to evaluate the students' knowledge, x^2 (1, $n = 124$) = 11.70, $p= <.001$, phi= .324 (*medium size effect*). Indeed, more females (76.6% or 49 out of 64) use CP than males (45% or 27 out of 60). Results also show a significant difference in the proportion of female (18.8% or 12 out of 64) and male (41.7% or 25 out of 60) teachers who are using 'lab work' (LW) to evaluate the students' knowledge, x^2 (1, $n = 124$) = 6.72, $p= .010$, phi= -.250 (*very small size effect*). For N-TAA, significant differences were found in the proportion of females (87.5% or 56 out of 64) and males (71.7% or 43 out of 60) who use "asking questions on previously taught content at the start of each class (QPC)" to evaluate the student's knowledge, x^2 (1, $n = 124$) =3.9, $p=.049$, phi= 0.197 (*small effect size*). No significant differences were found based on gender for all other N-TAA items.

E. What features do educators believe are effective in assessing how well students perform in the course?

This question focused on the key features teachers observe with students who perform well in the course (see Table II). A

TABLE II: Features for identifying students who perform well

Indicators	N	Mean	Std. Deviation
Attendance at classes/lectures	123	**3.34**	0.711
Asking questions during or out with the lectures (Asking_Q)	124	**3.48**	0.681
Answering questions during or out with the lectures (Ans_Q)	124	**3.44**	0.653
Help/answers provided to other students (Help_Peers)	123	3.25	0.742
Use of additional resources provided in the course (e.., office hours, revision lectures, additional tutorials, etc.) (Add_rec)	122	3.23	0.769
Timely submission of assignments (Timely_Sub)	124	3.15	0.772
High scores in summative assessment (High_Score)	120	3.28	0.688

one-way analysis of variance (ANOVA) revealed no significant differences in perceptions of the effectiveness of the features, F (2, 121) =2.02, p=.14 among teachers with varying years of experience between 'up to 10 years' ($n = 44$, $M = 3.2$, $SD = .46$), '11-15 years' ($n = 32$, M=3.3, SD=.4) and '16+ years' groups (n=48, M= 3.4, SD=.5). However, an independent-sample t-test revealed a significant difference in the effectiveness of feature scores between males (M=3.21, SD=.5) and females (M=3.4, SD=.4; t (121) =2.21, p=.028, two-tailed. The magnitude of the differences in the means (mean difference = .183, 95% CI [-.35, -.02]) was small (Cohen's d = -.4). All seven items in the scale have a mean of +3 (very effective, somewhat effective). Asking questions (Asking_Q), answering questions (Ans_Q), and attendance receive top means of effectiveness with 3.48, 3.44, and 3.34, respectively. Timely submission recorded the highest rating in terms of ineffectiveness (somewhat ineffective and very ineffective, $n = 16$).

V. Discussion

In this study, we investigated teachers' perceptions of the accuracy of traditional and non-traditional assessment activities they use to evaluate their students' learning status in an online learning environment. The study also investigated the features teachers believe are effective in identifying how well students perform in the course. Findings showed that despite having opportunities to use different ways of assessing student learning in an online environment [24] most teachers still prefer using traditional assessment activities such as final exams, individual assignments, and mid-term exams. This is also supported in a previous study by [25] which showed that exams are best at measuring learning compared to other activities. Another study also highlighted that summative assessments remain the best activities to evaluate student learning [26]. Despite the benefits of peer review in teaching and learning [27], in this study, peer review was one of the least used traditional assessment teaching activities and it was also not identified as an activity that can accurately evaluate students' knowledge. Indeed, several studies have shown the positive impact of peer review on student learning [28], [29], [30]. However, as other studies [31] have highlighted this reluctance to use peer review as a summative assessment could be because teachers lack faith in their ability to understand the guidelines, standards, and expectations for professional peer reviews of learning.

Asking questions on previous and current content is the most used set of non-traditional assessment activities, which teachers also believe to be the most accurate in helping them evaluate student knowledge. This agrees with [22] who also found that questions teachers ask can elicit prior knowledge and verify students' developing understanding. On the other hand, students asking and answering during the lecture were identified as the most effective features teachers observe that help them determine how well students perform in the course. Indeed, students who tend to ask or answer questions are more confident [32], and confidence correlates with high student performance and engagement [33].

Overall, there were no significant differences in perceptions of the accuracy of traditional and non-traditional assessment activities among teachers with varying years of experience (up to 10 years, 11-15 years, and 16+ years). However, teachers with online teaching experience of more than four years were more likely to use an electronic voting system (EVS), ask their students to record a short video (SVR) and use One-time paper asking students to record their thoughts (OTP) than those with less than four years online teaching experience. Some studies have shown that EVS [3] can have a positive impact on learning however this may require technical skills or training [34]. Moreover, other studies have shown that the longer the teaching experience, the more confident teachers are in using new techniques or teaching activities [19].

Overall, there were no significant gender differences in the perceptions of the accuracy of traditional and non-traditional assessment activities. However, female teachers tend to use class participation (TAA) and ask questions on previously taught content at the start of each class (N-TAA) more than their male counterparts, who instead use lab work the most as TAA. Indeed, some studies have shown that female teachers prefer more active learning activities than males [20].

The study asked teachers to rate the features they observe in students who perform well in the course (learning indicators). All the identified features were viewed as somewhat and very effective as learning indicators. This is not surprising. For example, some studies have highlighted attendance as critical in student progression and retention [21]. Furthermore, access to additional resources has been positively linked with access to extra material [35]. However, female teachers found these features more effective in helping them than males. As observed in previous studies, female teachers foster active learning activities, including encouraging students to ask and answer questions and helping their peers [20].

Findings showed that educators do not regard timely submission of assessments as an effective feature to inform whether a student performs well in the course. This disagrees with the literature [36], indicating that completing the assignments in a timely manner is a feature of successful students. While this

study uses a small sample size used, it lays the foundation for future work, which could involve a larger sample.

VI. Conclusions

In this paper, we investigated teachers' perceptions of the accuracy of traditional and non-traditional assessment activities they use to evaluate their students' learning status in an online learning environment. The study also investigated the features teachers believe effectively identify how well students perform in the course. Results showed that despite the preference for traditional assessment activities such as final exams, individual assignments, and mid-exams, non-traditional assessment activities could also accurately indicate students' level of understanding in online education. This study strongly indicates that non-assessment-related activities, such as asking questions, and discussion forums, can be effective learning indicators. This finding has implications for academic institutions and educators interested in supplementing traditional approaches to assessing student learning with non-traditional methods. This forms the basis for future work, in which the authors will implement the investigated learning activities to assess their effectiveness.

References

[1] D. Clow, "MOOCs and the funnel of participation," in *Proceedings of Learning Analytics and Knowledge (LAK)*, pp. 185–189, 2013.

[2] E. M. Onyema, E. C. Deborah, A. O. Alsayed, Q. Noorulhasan, and S. Sanober, "Online discussion forum as a tool for interactive learning and communication," *International Journal of Recent Technology and Engineering*, vol. 8, no. 4, pp. 4852–4859, 2019.

[3] A. I. Wang and R. Tahir, "The effect of using kahoot! for learning–a literature review," *Computers & Education*, vol. 149, p. 103818, 2020.

[4] E. M. Stacy and J. Cain, "Note-taking and handouts in the digital age," *American journal of pharmaceutical education*, vol. 79, no. 7, 2015.

[5] C. J. Brame, "Effective educational videos: Principles and guidelines for maximizing student learning from video content," *CBE—Life Sciences Education*, vol. 15, no. 4, p. es6, 2016.

[6] G. Norman, A. Neville, J. M. Blake, and B. Mueller, "Assessment steers learning down the right road: impact of progress testing on licensing examination performance," *Medical teacher*, vol. 32, no. 6, pp. 496–499, 2010.

[7] S. Pokhrel and R. Chhetri, "A literature review on impact of covid-19 pandemic on teaching and learning," *Higher education for the future*, vol. 8, no. 1, pp. 133–141, 2021.

[8] M. Sadeghi, "A shift from classroom to distance learning: Advantages and limitations," *International Journal of Research in English Education*, vol. 4, no. 1, pp. 80–88, 2019.

[9] I. A. Mastan, D. I. Sensuse, R. R. Suryono, and K. Kautsarina, "Evaluation of distance learning system (e-learning): a systematic literature review," *Jurnal Teknoinfo*, vol. 16, no. 1, pp. 132–137, 2022.

[10] W. Harlen, C. Gipps, P. Broadfoot, and D. Nuttall, "Assessment and the improvement of education," *The curriculum journal*, vol. 3, no. 3, pp. 215–230, 1992.

[11] K. E. Dunn and S. W. Mulvenon, "A critical review of research on formative assessments: The limited scientific evidence of the impact of formative assessments in education," *Practical Assessment, Research, and Evaluation*, vol. 14, no. 1, p. 7, 2009.

[12] J. B. Hansen, "Student performance and student growth as measures of success: An evaluator's perspective," *Annual Meeting of the American Educational Research Association*, 2000.

[13] D. Wiliam, "What is assessment for learning?," *Studies in educational evaluation*, vol. 37, no. 1, pp. 3–14, 2011.

[14] D. Woit and D. Mason, "Effectiveness of online assessment," in *Proceedings of the 34th SIGCSE technical symposium on Computer science education*, pp. 137–141, 2003.

[15] P. C. Abrami, "On the nature of support in computer-supported collaborative learning using gstudy," *Computers in Human Behavior*, vol. 26, no. 5, pp. 835–839, 2010.

[16] P. C. Abrami, R. M. Bernard, E. M. Bures, E. Borokhovski, and R. M. Tamim, "Interaction in distance education and online learning: Using evidence and theory to improve practice," *Journal of computing in higher education*, vol. 23, no. 2-3, pp. 82–103, 2011.

[17] S. N. Kew and Z. Tasir, "Learning analytics in online learning environment: A systematic review on the focuses and the types of student-related analytics data," *Technology, Knowledge and Learning*, pp. 1–23, 2022.

[18] N. Rahmah, A. Lestari, L. A. D. Musa, and H. Sugilar, "Quizizz online digital system assessment tools," in *2019 IEEE 5th International Conference on Wireless and Telematics (ICWT)*, pp. 1–4, IEEE, 2019.

[19] T. Dolighan and M. Owen, "Teacher efficacy for online teaching during the covid-19 pandemic," *Brock Education Journal*, vol. 30, no. 1, pp. 95–95, 2021.

[20] E. Sabbe and A. Aelterman, "Gender in teaching: A literature review," *Teachers and Teaching: theory and practice*, vol. 13, no. 5, pp. 521–538, 2007.

[21] A. Davis, "The correlation between attendance and achievement," *Foreword Summary Overview of Projects Teaching Fellowship Reports College of Applied Arts and Tourism*, vol. 1, no. 1, pp. 10–16, 2011.

[22] C. V. Schwarz, C. Passmore, and B. J. Reiser, *Helping students make sense of the world using next generation science and engineering practices*. NSTA Press, 2017.

[23] D. Bertram, "Likert scales," *Retrieved November*, vol. 2, no. 10, pp. 1–10, 2007.

[24] I. Yuhanna, A. Alexander, and A. Kachik, "Advantages and disadvantages of online learning," *Journal Educational Verkenning*, vol. 1, no. 2, pp. 13–19, 2020.

[25] B. L. Bartoszewski and R. A. Gurung, "Comparing the relationship of learning techniques and exam score.," *Scholarship of Teaching and Learning in Psychology*, vol. 1, no. 3, p. 219, 2015.

[26] F. J. Cilliers, L. W. Schuwirth, N. Herman, H. J. Adendorff, and C. P. van der Vleuten, "A model of the pre-assessment learning effects of summative assessment in medical education," *Advances in Health Sciences Education*, vol. 17, pp. 39–53, 2012.

[27] J. Sachs and M. Parsell, *Peer review of learning and teaching in higher education: International perspectives*, vol. 9. Springer Science & Business Media, 2013.

[28] S. Hopewell, G. S. Collins, I. Boutron, L.-M. Yu, J. Cook, M. Shanyinde, R. Wharton, L. Shamseer, and D. G. Altman, "Impact of peer review on reports of randomised trials published in open peer review journals: retrospective before and after study," *Bmj*, vol. 349, 2014.

[29] K. Lundstrom and W. Baker, "To give is better than to receive: The benefits of peer review to the reviewer's own writing," *Journal of second language writing*, vol. 18, no. 1, pp. 30–43, 2009.

[30] C.-C. Liu, K.-H. Lu, L. Y. Wu, and C.-C. Tsai, "The impact of peer review on creative self-efficacy and learning performance in web 2.0 learning activities," *Journal of Educational Technology & Society*, vol. 19, no. 2, pp. 286–297, 2016.

[31] R. R. Cavanagh, "Formative and summative evaluation in the faculty peer review of teaching," *Innovative higher education*, vol. 20, pp. 235–240, 1996.

[32] K. Kay and C. Shipman, "The confidence gap," *The Atlantic*, vol. 14, no. 1, pp. 1–18, 2014.

[33] J. Moneva and S. M. Tribunalo, "Students' level of self-confidence and performance tasks," *Asia Pacific Journal of Academic Research in Social Sciences*, vol. 5, no. 1, pp. 42–48, 2020.

[34] A. Comas-Quinn, "Learning to teach online or learning to become an online teacher: An exploration of teachers' experiences in a blended learning course," *ReCALL*, vol. 23, no. 3, pp. 218–232, 2011.

[35] S. W. Burleson, *Getting extraordinary things done: Perceptions of behaviors that superintendents use to obtain and manage additional resources in ways that result in improved student learning*. University of La Verne, 1998.

[36] J. W. You, "Identifying significant indicators using lms data to predict course achievement in online learning," *The Internet and Higher Education*, vol. 29, pp. 23–30, 2016.