An Investigation on Assessment Strategies, Student Engagement, and Retention for Large Cohorts Affected by COVID Learning Disruptions

Jeannette Chin

School of Computing Sciences, University of East Anglia, Norfolk, UK
j.chin@uea.ac.uk

Abstract. This paper reports the results of an investigation of assessment strategies for student learning, engagement, and retention of an undergraduate year 1 module in the Computing Science discipline at the University of East Anglia, UK. In this study, three different assessment methods were considered, one for each year, over a three-year period (2020-21 to 2022-23). The study period coincided with the COVID pandemic where the cohorts had their secondary school learning disrupted one way or another, prior to embarking on their university career. The assessment methods investigated did not cover all of the learning objectives, however the learning objectives assessed were comparable with one another. The results show that the presentation and in-class test assessment methods achieved normal distributions of marks, while the marks for the practice-based portfolio assessment were negatively skewed, suggesting the nature of the assessment requires more balancing tasks. Further, student attendance and submission rate were found to have been influenced by the assessment type students had to undertake. Cohorts who undertook the practice-based portfolio assessment had better student engagement and submission rate, at 73% and 91.92% respectively. Finally, learning disruption caused by the COVID pandemic was found to be correlated with student retention, where cohorts whose grades were determined solely by their teacher prior to attending university had a 24% higher chance of withdrawing from the course or transferring to a different course compared to those whose grades were determined by exams.

Keywords: Assessment Method, Student Engagement, Student Retention, Large Cohort, Computing Science, COVID Pandemic

I. Introduction

Student intakes for the Computing Science discipline have increased steadily in recent years due to industry demands for good computing skills. The trend is largely driven by the advancement of computers and digital technologies, coupled with pervasively available and increasingly reliable global network connectivity, paving new ways for modern data acquisition in many sectors. Ultimately the trend leads to the rise of Artificial Intelligence, pushed by companies in the high-tech sector, creating the perfect timing for the discipline to take the centre stage of this digital phenomenon. Many

universities, particularly in the UK, have experienced a surge in demand for places in Undergraduate admission in this discipline.

It seems that a long time has passed since the world recovered from the COVID-19 (COVID) pandemic even though the disease is still here with us. During the pandemic, like in other sectors, education had been severely disrupted in the UK, with classes either cancelled or moved online. The disruption had led to crucial exams for schools being cancelled for two consecutive years whilst the UK government employed strategies to determine the pupils' grades instead. For example, in the early days of the pandemic (academic year 2019-20), a strategy that saw the pupils' grades determined by their teachers and moderated by an algorithm was used. While the use of an algorithm was designed to reflect upon the grades more accurately, it introduced unwanted bias for pupils in certain demographics, leading to non-satisfactory results. Consequently, a different strategy was needed the following year (academic year 2020-21), which saw the pupils' grades determined only by their teachers. Learning disruption eased off when the pandemic was under control in 2022, which saw crucial exams resuming in the same year. However, given pupils all had their learning negatively affected by the pandemic, a different strategy was needed to gently ease them back to the "normal" education system. This meant that year, additional information about the content of the exams were supplied to teachers and pupils in advance to help them prepare. Learning disruption caused by COVID brought huge challenges to the higher education system in the UK simply because it relies on exam grades to evaluate the suitability of students.

The paper presents the results of an investigation of three assessment methods. The study was carried out on a Year 1 computing module with a large cohort of students, spanning a three-year period (2020-21 to 2022-23). The cohort consisted of students taking courses on computing pathways and from other schools within the faculty. Entry grades were the same for all students in the entire study period. This means all students obtained the entry grades needed to secure a place in their respective course, but how their grades were determined was very different from year to year, as explained earlier. The study explored three assessment methods, one for each year. Weekly teaching was delivered via lectures and laboratory sessions. However, the mode of delivery for each cohort was adjusted according to COVID restrictions at that time. For example, online learning for the 2020-21 cohort (because of the lockdown), a hybrid mode of learning for the 2021-22 cohort (due to COVID restrictions such as number of people allowed to be in a room), and in-person learning for the 2022-23 cohort (all restrictions lifted).

The paper is structured as follows: Section II describes background information, assessment strategies and pedagogy, Section III discusses results, student engagement and retention, and Section IV concludes the work.

II. Background and Pedagogy

The module concerned is a core year 1 module for BSc Computing Science at the University of East Anglia, UK. It is also an optional module for other courses within the Faculty of Science. The student cohorts in this investigation period were: 2020-2021 - 232 students, 2021-2022 - 193 students, and 2022-2023 - 223 students (Table 1).

The module is about Web Programming, a complex concept. This means, at a minimum, the learning involves mastering different types of techniques from different scripting or declarative languages, each having its own syntax and usage. It also involves the learning and understanding of the Internet, data communication methods and the use of other technologies / tools (e.g. a browser or console) for viewing programming results. The module is a semester-based module, running for 12 weeks.

Table 1. Assessment method used for each cohort.

Cohort	Size	Assessment	Learning Mode
2020-21	232	Portfolio	Online
2021-22	193	Presentation	Hybrid
2022-23	223	In-Class Test	In-Person

Robins et al [9] examined multiple programming pedagogy for new learners, including the use of pair-programming and visualisation tools. In a review [10] they suggested that for delivering effective learning, simply illustrating the concepts is insufficient. Effort should be spent on teaching the process of programming instead. One of the effective leaning strategies is to adopt "trial and error" approach using scaffolding [13] or live coding [11]. Scaffolding is a teaching strategy where support is provided to students initially to guide them solving problem. However, this support should be gradually removed as students gain more experience and knowledge to encourage independent learning. Live coding is an effective teaching method where coding techniques and solutions are demonstrated "live" together with the students to enhance their understanding. Here, for this study, multiple strategies were adopted in the teaching delivery. For example, the module teaching plan included a series of lectures illustrating the concepts and teaching the process of programming using a combination of talks, short videos, and live coding. This was followed by hands-on practical sessions where full support was provided to guide the students for a duration of six weeks. The full support would then be replaced by partial support until the end of term. Table 1 above shows the assessment methods used for each of the cohort.

A. Assessment Strategies

Over the past few decades various assessment strategies have been developed [6][8] to facilitate active learning as an alternative to conventional assessment methods such as exams or essays. This could be seen as a direct result of several critiques arguing against the over-reliance on conventional assessment methods. The arguments were centred round the rigidity and inflexibility of such assessment methods, making it unsuitable for certain learning, or worse, causing anxiety to students due to the absence of instruments to promote their mental well-being [1][4][9]. In addition, conventional assessment methods were often perceived as a measure of ability to gather facts or details and present the information in a clear and concise way [8]; it does not require students to channel their efforts into trying to understand the material that they study [12]. Further,

it has been well-documented that students adopt either surface-level or deep-level approach for their study [3][4][7][16]. Their decision, by and large, is influenced by the types of assessment they had to undertake [2][14]. Although students expressed positive attitudes towards multiple-choice assessments [16], these assessments are found to have encouraged surface-level approach to learning, to which it does not help students to retain their knowledge in the long-term [8]. This study employed two alternative assessment methods – practice-based portfolios and oral presentations, along with a time-constrained in-class test.

B. Study Design and Procedure

The primary goal of this study is to investigate which assessment method is effective and efficient against student learning, engagement, and retention. Given the cohorts investigated all had their secondary school learning disrupted by the COVID pandemic, the study seeks to explore the impact on learning, engagement and retention as well. Three assessment methods were selected to be evaluated in this study (Table 1):

1. A Practice-based Portfolio
   The portfolio assessment was perceived as being an assessment tool helping students to learn better and retain information longer, thus a preferred choice of alternative assessment for students [15]. This can be attributed to the fact that the portfolio assessment requires students to work with the material over an extended period of time. Slater [15] suggested that by having students continuingly working with the material, the portfolio encourages students to think about the concepts creatively and apply them in a variety of ways. Moreover, students also have been found to have enjoyed the time they spent on creating portfolios, which was said to be a rewarding and meaningful experience for them [15]. For this study, the 2020-21 cohort took this assessment. A practice-based portfolio was carefully designed and aligned with weekly learning objectives. A clear marking scheme was developed in the process. The portfolio contained five weekly lab coding exercises linked to the topic covered for that week. For example, HTML was covered in week 1, students were required to complete the HTML coding exercises that week and so on. Each coding exercise had a difficulty level ranging from super easy, moderate easy, to difficult (require more effort). Online teaching was used due to COVID lockdown restrictions at the time. Students were required to work individually to complete the portfolio. The assessment was due mid-term.
2. An Oral Presentation Assignment
   Clear assessment criteria and standards can help students to understand what is expected of them and to set goals for their own learning [5]. For this study, the 2021-22 cohort took this assessment. An oral presentation assignment was carefully developed and aligned with weekly learning objectives. Based on a specific scenario, the assignment evaluated five weeks of learning outcomes where students were required to design and code a web user interface and present their solution orally. Each student was given 10 minutes to present their solution. The assignment goals, which students should aim to achieve, were clearly highlighted in the assignment brief. A grid of marking criteria was also developed and provided in

the brief where students could easily refer to. Hybrid teaching was used - online (lectures) and in-person (laboratory sessions) due to COVID restrictions at that time. Students were required to work individually for this assessment. The assessment was completed mid-term.

3. A Time-Constrained In-Class Test
   The time-constrained in-class test can be considered as one of the conventional assessment methods. A study [13] found that students preferred multiple-choice tests over essay writing. This preference was said to have been drawn by the perception of such tests being easier to prepare for, easier to take, and may produce higher relative scores. However, the study also found that multiple-choice tests only encourage students to adopt surface-level approach to learning [8]. For this study, the 2022-23 cohort took this assessment. A set of online summative and formative in-class tests were carefully designed and developed to cover weekly learning objectives. Upon reflecting on the literature findings, both summative and formative in-class tests included a combination of question types, including multiple-choice, "fill in the blanks", and "short answer" types of questions. A marking scheme and feedback mechanism were developed and implemented in an online system in such a way that marking could be done automatically by the system. In-person teaching was used due to COVID restrictions having been lifted. Students were required to work individually for this assessment. The summative test was scheduled to be taken outside the teaching semester, but students were provided with weekly formative tests to help them to prepare.

III. Results and Discussions

The results of each year are presented in this section along with discussions. Non-submission cases are excluded in these results but are discussed in Student Engagement section.

A. Results

- Cohort 2020-21 – Practice-based Portfolio
  206 out of 232 students (88.79%) undertook this assessment. The average mark is 79.01 (STD=15.437, N=206), Fig. 1.
  With an average mark of 79.01 (Fig. 1), the result shows that a lot of students did exceptionally well for this assessment while some did not. For those who had submitted, 1% (sum=2) of the students failed this assessment (pass mark=40). Further test for Normality were conducted using SPSS.

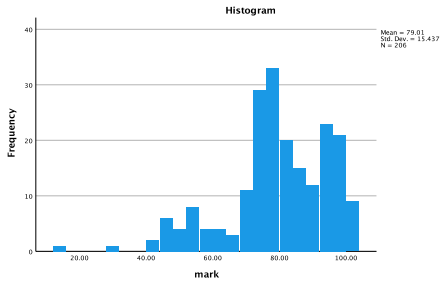

Fig. 1. The mark distribution for cohort 2020-21

Given the sample size of this cohort is not large, Shapiro-Wilk test was performed with a confidence interval set at 95%. The results showed the distribution of marks significantly departed from normality ($W = 0.922$, $p < 0.01$) (Fig. 2). This can be also observed from the histogram that the data are far from normally distributed (Fig. 1).

Shapiro-Wilk Test

	Statistic W	df	Sig.
mark	.922	206	<.001

Fig. 2. Test of Normality for cohort 2020-21

- Cohort 2021-22 – Oral Presentation
  165 out of 193 students (85.49%) undertook this assessment. The average mark is 64.01 (STD=11.432, N=165), Fig. 3.

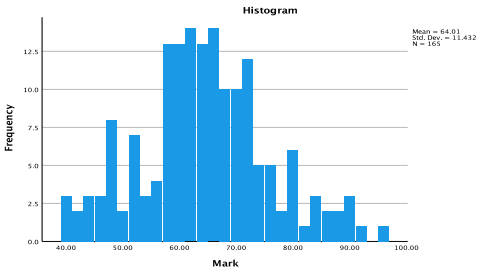

Fig. 3. The mark distribution for cohort 2021-22

A minimum mark of 40 and maximum mark of 95 was observed (Fig. 3). Further test for Normality were conducted using SPSS. Again, given the sample size of this cohort is small, Shapiro-Wilk test was performed with a confidence interval set at 95%. Here, again, the results show evidence of marks in

non-normality distribution (W = 0.986, p > 0.05) (Fig. 4). Based on these results and after visual examination of the histogram (Fig. 3), the mean with standard deviation were used to summarise the data. For those who had submitted none failed this assessment (pass mark=40).

Shapiro-Wilk Test

	Statistic W	df	Sig.
Mark	.986	165	.099

Fig. 4. Test of Normality for cohort 2021-22

- Cohort 2022-23 – Time-Constrained In-Class Test
  205 out of 223 students (91.92%) undertook this assessment. The average mark is 61.96 (STD=13.67, N=205), Fig. 5.

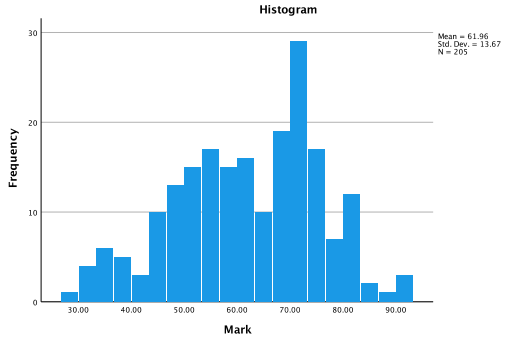

Fig. 5. The results for cohort 2022-23

A minimum mark of 29.4 and maximum mark of 92.04 were observed (Fig. 5). Further test for Normality were conducted using SPSS. Given the sample size of this cohort is also small, Shapiro-Wilk test was performed with a confidence interval set at 95%. Again, the results show evidence of data in non-normality distribution (W = 0.981, p > 0.05) (Fig. 6). Based on these results and after visual examination of the histogram (Fig. 5), the mean with standard deviation were used to summarise the data. For those who took the test, 8% (sum=16) failed this assessment (pass mark=40).

Shapiro-Wilk Test

	Statistic W	df	Sig.
Mark	.981	205	.073

Fig. 6. Test of Normality for cohort 2022-23

From the results it can be observed that while the minimum mark for the first cohort is the lowest (min=12), this cohort is the only cohort that achieved the highest possible maximum mark (max=100). To examine the data further, a Quartile Index (QI) for each cohort was generated, shown in Table 2. From the QI, it has been observed that the first cohort did exceptionally well in the practice-based portfolio assessment, with quartile indexes of 72.5, 79, 92.5 and 100. 28.49% of students' scores were less than 72.5, while large majority (71.51%) of students' scores were above 72.5 (Fig.7). The disparity between extremely high and low marks could be attributed to the nature of the assessment where tasks were mostly descriptive even for some of the more challenging questions. In addition, feedback students received during laboratory sessions had helped them complete the tasks and getting high scores. The difference between each QI for the remaining two cohorts is between 2 to 5 and the indexes lie within the expected range. The 2021-22 cohort has roughly an equal distribution of marks for each quartile, 26.24% of students' scores were below 58, 22.70% students' scores were between 58 and 64, and 51.06% of students' scores were above 71(Fig. 7), a 20.45% drop in marks greater than 71 compared with the previous cohort. The results could be attributed to the nature of the presentation assessment, which required students to apply the knowledge they gained and present their solution orally. This cohort had the highest non-submission rate which is discussed in Student Engagement section (III.2). It has been observed that the number of students scoring above 72 had dropped further for the 2022-23 cohort with 49.5% of students' scores were above 72.24 (22% drop compared with the first cohort). 25.5% of students' scores were between 63.24 and 72.24, and 25% scored below 63.24 (Fig.7). These results were expected given the nature of the assessment (a hybrid form of conventional assessment method), which was time-based and closed-book test.

B. Student Engagement

Student Engagement (SE) is an ambiguous term, which often entails confusion with people having different views and understanding about the meaning and what should be involved. Thus, SE can have different meanings depending on the stakeholder. For this study, the following simple metrics were used to measure and evaluate SE:

1. Were the students attending their classes regularly?
2. What is the submission rate?
3. Have the students passed the module?
4. Did the students score well?

Table 2. Quartile Index

Cohort	Min	Q1	Q2	Q3	Max
2020-21	12	72.5	79	92.5	100
2021-22	40	58	64	71	95
2022-23	29.4	52.12	63.24	72.24	92.04

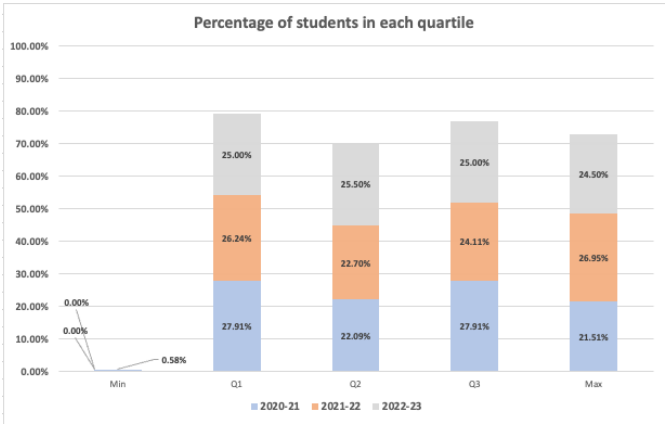

Fig. 7. Percentage of students in each quartile

The assumption here is that if the curriculum (pedagogy) was designed well and delivered smoothly, then in theory, students would have very few reasons not to attend their classes and learn. For clarity, Extenuation Circumstances and self-study time are excluded in this study.

An investigation into SE covering the same three-year period for the same module was conducted. To answer question (1), a digital QR recording system was used weekly to capture Student Attendance. Given the module is a hands-on module, the focus is on laboratory attendance. Laboratory attendance data for all three cohorts was collected from the university IT system. An Average Attendance Rate (AAR) was calculated for each year, shows in Table 3. It has been observed that all cohorts had an AAR over 60%. More than two third of the students of first and last cohorts (AAR=73%) attended their lab sessions fairly regularly. This result is quite interesting as the delivery for 2020-21 cohort was online due to COVID lockdown that year, while for 2022-23 cohort, the delivery was in-person as COVID restrictions had been lifted. These two cohorts were also required to complete their weekly practice-based laboratory exercises as part of their summative assessment, which may have contributed to high attendance rates. For the 2021-22 cohort, although students were encouraged to complete their weekly laboratory exercises, they were assessed on a separate assignment via an oral presentation. This could be the reason why the attendance for this cohort was comparatively poor since students might be adopting a surface-level approach to learning, focusing on the summative oral presentation, and losing sight of formative learning. The last cohort (2022-23) was encouraged to take their weekly formative tests at a time that suited them but still maintained an excellent AAR. This is due to the fact that, like the first cohort, this cohort were required to complete their weekly practice-based laboratory exercises as part of their second group-based assignment, which is not reported in this paper. The AAR rates suggest that practice-based laboratory exercises accumulated to summative assessment had a large influence over student attendance and learning, evident in the student performance presented in Results section above.

Table 3. Student Attendance Rate

Cohort	Teaching Mode	AAR	Method
2020-21	Online only	73%	Portfolio
2021-22	Hybrid	61%	Presentation
2022-23	In-Person	73%	In-Class Test

To answer question (2), cohort data was gathered and compared with submission records, shown in Table 4. The Submission Rate (SR) is calculated based on the number of students who submitted the assignment for this study. The results show that the cohort who took the time-constrained in-class test assessment had the best submission rate at 91.92%, followed by the cohort who undertook the practice-based portfolio at 88.79%, followed by the cohort who undertook the presentation at 85.49%. Cohorts that had a non-submission rate greater than 10% was a cause of concern. This is discussed in Student Retention section (III.3) below.

Table 4. Submission Rate

Cohort	Size	Non-Submission	SR
2020-21	212	11.21%	88.79%
2021-22	193	14.51%	85.49%
2022-23	223	8.08%	91.92%

C. Student Retention

Retention is an area can be used to measure student learning and leading to student success. In 2021, 72% of undergraduates in UK universities completed their course, which was significantly higher compared to many countries. To understand whether COVID learning disruptions have had an influence over the retention of the course, further data analysis was carried out. Given the module is year 1 module, the study only considered those who either withdrew or transferred out of the course after the first year of study. Because of the unprecedented time caused by COVID pandemic, where UK national exams had to be cancelled, the data is compared with how the students' grades were determined for that particular year (Table 5). The results suggest that teacher-assessed grades did have an influence over student retention for first year of study. For example, more than 12% of the students in cohort 2020-21, whose grades were determined by their teacher and moderated by an algorithm, either withdrew from the course or transferred to another course after the first year of study. Nearly a quarter (24.6%) of the students in the 2021-22 cohort, whose grades were determined solely by their teachers, found themselves in similar situations, it is 3-fold higher compare to 2022-23 cohort. It is acknowledged that there are limitations in this analysis due the size of dataset. Further research into this matter would be useful for in-depth insights. Nevertheless, the limited data provides a snapshot on Student Retention directly impacted by the COVID pandemic on a course level.

Table 5. Student Retention Rate after first year of study

Cohort	Withdrawn or Transferred	Admission Grades
2020-21	12.8%	Exams cancelled. Teacher and algorithm assessed
2021-22	24.6%	Exams cancelled. Teacher assessed only.
2022-23	8.4%	Exams resumed.

IV. Conclusion

The paper investigated three assessment methods for a year 1 module in an undergraduate Computing Science course at UEA over a three-year period. None of these assessment methods met all the learning outcomes of the module concerned but the learning outcomes assessed were comparable with one another. All cohorts had their high school learning disrupted by COVID pandemic prior to embarking their university career. The study also drew a comparison between the mode of teaching delivery and cohorts learning, due to COVID restrictions at that time. Further, the study also investigated student engagement based on a set of simple metrics. The study found that the oral presentation and time-constrained in-class test assessment methods achieved a normal distribution of marks while the practice-based portfolio assessment had a negatively skewed mark distribution, reviewing a large disparity of marks. This result was mainly attributed to the nature of the assessment in question, which was descriptive in nature, and could be made to challenge students' own initiative more. It was acknowledged that support and feedback provided had helped with the results. Further, the learning mode (online only) caused by the COVID lockdown at that time had an impact on student learning. Students who found online learning challenging did not do very well while students who managed to cope did the opposite. The practice-based portfolio accumulating to summative assessment is found to have a great influence on student attendance and submission rate. Cohorts who undertook this assessment had the highest attendance and submission rates, at 73% and 91.92% respectively. The time-constrained in-class test assessment method is found to be most efficient and effective as (1) it encourages a deep-level approach to learning (as it included not just multiple-choice questions but also other question types assessing the application of the concepts), (2) achieves a good normal mark distribution reflecting on students' ability, (3) requires fewer human hours for marking (as the majority of the questions can be marked automatically by an online system), and (4) leveraging technology, consistent feedback can be provided with minimum effort. Student retention is also found to had been influenced by how their secondary school grades were determined. Students whose grades were determined solely by their secondary school teachers had a 24% higher chance to either withdraw or transfer after their first year of study. It has been acknowledged that there are limitations in this study due to the size of the dataset. Further research would be needed to provide greater insights on the learning impact caused by the COVID pandemic.

Acknowledgements
A special thanks to Chris Harfield at Learning and Teaching Service and Charlotte Caws at Data Quality & Standards Service at UEA for their diligent data collection work.

References

1. S. Cassidy and P. Eachus, "Learning style, academic belief systems, self report student proficiency and academic achievement in higher education," Educational Psychology 20: 307–322., 2000, doi:10.1080/713663740.
2. N.J. Entwistle, and A.C. Entwistle, "Revision and the experience of understanding," In The experience of learning, ed. F. Marton, D.J. Hounsell, and N.J. Entwistle, 145–58. Edinburgh: Scottish Academic Press, 1997
3. N. Entwistle, and P. Ramsden, "Understanding student learning (Routledge revivals)," Routledge; 20 Aug 2015.
4. G. Gibbs, and T. Habeshaw,"Preparing to teach," Bristol: Technical and Educational Services, 1989
5. J. Hattie, and H. Timperley, "The Power of Feedback," Review of Educational Research 77 (1): 81–112, 2007
6. C.M. Kell, and R.W.M. van Deursen. "Curricular influences on academic belief systems," Learning in Health and Social Care 1, no 2: 86–93, 2002
7. K.L. Krause, and H. Coates, "Students' Engagement in First-Year University," Assessment & Evaluation in Higher Education, 33, 493-505, 2008, doi:10.1080/02602930701698892
8. K. Struyven, F. Dochy, and S. Janssens, "Students' perceptions about evaluation and assessment in higher education: a review," Assessment & Evaluation in Higher Education, 30:4, 325-341, 2005, doi: 10.1080/02602930500099102
9. P. Race, "The lecturer's toolkit: A practical guide to learning," teaching and assessment. 3rd ed. London: Routledge, 2006
10. A. Robins, J. Rountree, and N. Rountree, "Learning and teaching programming: A literature review," Computer Science Education, 13(2):137–172, 2003.
11. M.J. Rubin. "The effectiveness of live-coding to teach introductory programming," In Proceeding of the 44th ACM technical symposium on Computer science education (SIGCSE '13). Association for Computing Machinery, New York, NY, USA, 651–656, 2013, doi:https://doi.org/10.1145/2445196.2445388
12. K. Sambell, L. McDowell, and S. Brown, "'But is it fair?': an exploratory study of student perceptions of the consequential validity of assessment," Studies in Educational Evaluation, 23(4), 349–371, 1997
13. A. Sanwar, "Effective teaching pedagogies for undergraduate computer science," Mathematics and Computer Education, 39(3):243–257, 2005.
14. K.M. Scouller, and M. Prosser, "Students' experiences in studying for multiple choice question examinations," Studies in Higher Education 19: 267–79, 1994.
15. T.F. Slater,"Portfolio assessment strategies for grading first-year university physics students in the USA," Physics Education, 31(5), 329–333, 1996
16. E.R. Traub, and K. MacRury, "Multiple-choice vs. free response in the testing of scholastic achievement," in: K. Ingenkamp & R. S. Jager (Eds) Test und tends 8: jahrbuch der pädagogischen diagnostik (Weinheim und Base, Beltz Verlag), 128–159, 1990