



AUTONOMOUS GRADER – A NEW DIGITAL SUMMATIVE ASSESSMENT TOOL

Mladen Vidović*,
Radmila Suzić,
Ivan Radosavljević,
Nebojša Nešić

Singidunum University,
Belgrade, Serbia

Abstract:

Assessment is an essential part of both traditional and online education. Regardless of the learning environment, assessment can be formative and summative, each with advantages and disadvantages in the digital era. In order to harness its possibilities in an online environment, educators must overcome certain challenges that technology-based assessment imposes. This paper includes an overview of the literature that covers tools that can be used as digital alternatives to traditional pen and paper tests. In order to maintain the positive learning environment in online education, assessment tools had to be adapted to the new learning environment in order to be accessible remotely to a large number of students. The paper discusses several advantages of digital assessment tools in light of their credibility, validity, and reliability, especially in case of summative digital tools, which includes the solution presented in this paper, the Autonomous grader. This tool, created with the purpose of promoting the use of technology in teaching, improves the testing system by offering effective, precise and unbiased grading of each individual assignment, reducing the time spent on grading. The proposed solution is a test environment utilizing Selenium tests. The solution includes a language for defining tasks and evaluation criteria used to generate individual Selenium tests. The result of the automatic evaluation process is a human-readable report containing the results of individual assignments. We conclude that this type of tool can be used effectively and can be a valuable asset to assessment in the digital age.

Keywords:

Summative assessment, online education, automated grading.

INTRODUCTION

Due to the abrupt change in the education system caused by the pandemic of COVID-19, the traditional forms of evaluation ceased to be used. Digital alternatives had to be found for both formative and summative assessment, which are considered to be two most occurring forms of assessment in any learning environment [1]. While formative assessment focuses on observing students' learning process, alternation, modification and improvement of their learning [1], summative assessment does the opposite. Summative assessment tools are standardized tests typically given upon completion of a course and are often seen as a quantitative appraisal of specific learning outcomes [1, 2].

Correspondence:

Mladen Vidović

e-mail:

mvidovic@singidunum.ac.rs



With the full change of focus from traditional pen and paper tests and in-person observations, both summative and formative assessments suffered great changes. Seeing students on devices, teachers started combining technology with assessment forms because they noticed the advantages of adapting it to the new learning environment. Using digital forms of formative and summative assessment has proven to provide faster and almost immediate feedback, allowing teachers more time for rapport with students. In the light of capabilities of learning in an online environment this paper will present the process of creation of a new summative assessment tool – Autonomous grader. This tool was created with the purpose to evaluate any structured document. The validation process is demonstrated on an end of the course summative assessment for undergraduate students at Singidunum University. Their final exam in the course focused on the utilization of modern technologies in language learning. It was based on the creation of an online summative assessment tool using already existing online learning platforms such as Google Classroom, Microsoft Teams, and ClassDojo.

2. THEORETICAL BACKGROUND

In recent times, there has been a widespread use of online tests in assessing students' knowledge. The increased need for new forms of assessment was derived out of a need to prepare students to be "work ready", competent and capable to meet the digital age [3]. Though many might have had doubts about the use of digital assessment tools, numerous benefits can be pointed out, not only concerning their efficiency, but pedagogical reasons as well. Digital tests are highly efficient due to automated grading and the possibility for immediate feedback, especially with large groups of students or in the case of different topics that need to be covered in one test of limited duration [4, 5]. When it comes to pedagogical reasons, digital assessment tools still struggle with traditionally oriented teachers, who doubt their validity, reliability and fairness [1, 6], which are the commonest reasons why they hesitate to implement these forms of testing in regular online or blended learning. However, the effectiveness of online tests should not be doubted if they are implemented in the context of an overall learning experience, and additionally supported by other forms of assessment [6]. Moreover, one should not forget the impact of individual and immediate feedback they provide [7, 8], which is likely to help students learning [9].

Digital assessment tools have been increasingly used in an online learning environment caused by the emergence of the world pandemic, in both formative and summative forms of assessment. Multiple choice questions, fill in exercises, online forms and quizzes are frequently used tools for summative assessment while discussion boards, reflections and small group discussions, digital portfolios are used for formative assessment [1, 10]. Since this paper focuses on the development of a new summative assessment tool, only forms focused on summative digital assessment and their characteristics shall be addressed in the following text.

One of the major benefits of summative digital assessment tools is their ability to provide automated answers and immediate scoring of evaluated data, which reduces the possibility for mistakes due to human factor. Digital scoring is effective and precise and the time invested in grading each individual paper is significantly reduced [2]. With the transition to online learning, many teachers decided to transform paper-based multiple-choice tests to digital format. It wasn't long before they realised the benefits of this format, because it allowed them to enrich their testing system with different media such as video, audio and other multimedia content.

Another advantage of digital assessment tools is the ease of their use. Once mastered, they are easily administrated to large groups, and the automated scoring process speeds up the administration of grades and its distribution to students [11]. Due to the minimization of human-made mistakes, there is an increase in efficiency which leads to their standardization. Another benefit refers to administration of digital tests and teacher's dialect, which is a significant feature concerning phonological awareness and listening comprehension [2]. One should not forget that technology-based assessment tools support educators in case of instructional modifications on an individual level which is necessary when dealing with students of different achievement profiles [12].

Digital assessment systems are surely the future that will cleverly encompass needs and perceptions of both teachers and students [13]. Teachers now have the task to make an effort to successfully integrate new technologies in their classrooms. Further advantages of technology based summative assessment tools are yet to be discovered, because with their increased use in an everyday classroom situation the potential for their advancement is inevitable.



Digital assessment systems are surely the future that will unavoidably encompass needs and perceptions of both teachers and students.

3. IMPLEMENTATION

The solution proposed in this paper is implemented as a sequence of Selenium tests, mimicking the usual workflow employed by teachers during the manual evaluation of students' assignments. The assignments are supplied by the teachers in the form of a structured document containing tasks. Each task is described using natural language, as to be delivered to students, a series of steps necessary to extract students' answers for a given task, and conditions that dictate when a task is considered successfully solved. Each condition corresponds to a single task and contains a template of an expected answer, as well as the score attained for successfully satisfying a condition. The template allows for a more flexible validation as opposed to the validation using strict equality with a specific value.

An example of a possible teacher-submitted task description is shown in Listing 1. Each variable defined in a GET clause can be used in the following get clauses and conditions. Variables defined in the GET clauses must be unique. Each variable can be assigned a set of values by using FROM SUCH THAT expressions. The FROM expression extracts values either from previously defined variables, constants, or implicitly defined variables such as the webpage variable. The SUCH THAT part of the FROM expression is optional and can be used to acquire a subset of values that match a specific criterion. Conditions are given as [VERBS] variable COMPARATOR [VERBS] value|variable SCORES score_value expressions where VERBS is an optional part of the expression consisting of a sequence of verbs acting on values or variables to extract additional data. The SCORES part of the expression is used to express the score a student is given if a condition is satisfied. The steps and conditions, given in a previously defined manner, of a task description, are translated into concrete Selenium tests. An example of the generated Selenium test for the task description shown in Listing 1 is shown in Listing 2.

```
Task: Create at least one mandatory multiple-choice question worth at least 5 points
Steps:
    Get quiz by URL url
    GET questions FROM quiz SUCH THAT type=multiple-choice AND required=true AND value>=5
conditions:
    Count questions > 0 SCORES 1
```

Listing 1 – An example of a simple task description.

```
@task("Create at least one mandatory multiple-choice question worth at least 5 points")
def task_1():
    score = 0
    Driver.get("url")
    questions = driver.find_elements(By.XPATH, "//question[@type='multiple-choice' and @
    required='true' and @value>=5]")
    if(len(questions) > 0):
        score += 1
    return score
```

Listing 2 – An example of a generated Selenium test.



Once a teacher collects all the students' works, they can initiate the automatic evaluation process. During the evaluation, an evaluation model is created for each submitted work. This model contains the total score per student, as well as triplets, one for each task, consisting of the student's extracted answer, the conditions the answer was required to meet, and the evaluated score. An example of a generated model is shown in Listing 3.

A human-readable report is also generated based on the models acquired by evaluating the students' work. The report is presented as an HTML page, shown in Figure 1, containing a table listing all the students and their scores, as well as links to individual models, with examples shown in Figure 2.

```
{
  "total_score":69,
  "Create at least one mandatory multiple-choice question worth at least 5 points" :
{"expected": ["count questions > 0"] , "submitted": ["count questions = 3"], "score": 1},
  ...
}
```

Listing 3 – An example of a resulting report model.

Assignment results

No.	Student ID	Total score	
1.	XXXXXXXXXX	2	View details
2.	XXXXXXXXXX	3	View details
3.	XXXXXXXXXX	0	View details
4.	XXXXXXXXXX	1	View details
5.	XXXXXXXXXX	1	View details
6.	XXXXXXXXXX	3	View details
7.	XXXXXXXXXX	0	View details
8.	XXXXXXXXXX	3	View details

Figure 1 – A table showing students' IDs and their scores for a given assignment.

Assignment overview

Student: XXXXXXXX

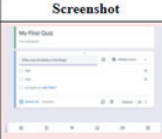
No.	Task	Score	Screenshot
1.	Create at least one mandatory multiple-choice question worth at least 5 points	0	
2.	Change the classroom title to contain between 10 and 20 characters and the term "ICT"	1	
3.	Add a topic with a link to a quizz to a classroom	1	

Figure 2 – An overview of a single submitted assignment. Pink rows indicate incorrectly completed tasks.



4. CONCLUSION

Making technology-based summative assessment tools part of regular educational practice has the potential to enhance the assessment process by increasing its credibility, validity, and reliability. However, numerous challenges and issues are commonly encountered when introducing innovative technologies in education. Those challenges include cheating as one of many practices students engage in when exposed to online or digital testing tools [14, 15]. In case of being assessed remotely, as it was the case during pandemics, students often took digital test as open-book tests, using numerous techniques such as multiple computers, online search engines, communication platforms and even helpful and knowledgeable friends [15] logging in from different devices under same name to help their classmates. However, many authors claim that the problem of cheating in using digital tools is not a big concern, since the same problem often occurs in face-to-face testing as well [16]. As in the case of taking regular pen-to-paper tests in the traditional classroom environment, teachers always have the opportunity to involve additional methods of testing their students' knowledge, such as an oral confirmation of the acquired material. Another drawback of the transition to digital assessment tools is the lack of technology requirements. Although it is a common occurrence today that every household has a computer and Internet connection, one must not forget multiple-children families where two or more children must share the same device. Even if a device is available to each student, not all students are experienced computer users and many will need further help or training in order to familiarise themselves with this type of assessment [17, 18, 19].

However, the benefits of using digital assessment tools outweigh the challenges, since they are proven to be extremely accurate and fast method of students' evaluation in comparison to traditional paper-based tests. The benefits of using these forms of assessments are numerous and present in different domains; these benefits are utilized by teachers and students in educational institutions to achieve educational aims [20].

Benefits for students are seen in friendly interfaces and gamification of assessment materials, and the possibility for numerous simulations of testing environment through repetition of the material that needed to be worked on [21]. Digital assessment tools, especially summative ones, offer immediate feedback that can further enhance students' learning, especially if students are located in remote areas or are employed.

Benefits for teachers are numerous, starting from the fact that digital assessment tools are less time consuming, unbiased, precise and accurate, with immediate feedback for each assignment. It also enables teachers to monitor students' progress analysing their performance across many assignments [22].

Digital assessment tools support educational goals of an institution showing that the institution is willing to adapt to the challenges of the new educational environment in order to support learning and teaching in the digital era.

5. REFERENCES

- [1] D. Perera-Diltz and J. Moe, "Formative and Summative Assessment in Online Education," *Journal of Research in Innovative Teaching*, vol. 7, 2014.
- [2] M. M. Neumann, J. L. Anthony, N. A. Erazo and D. L. Neumann, "Assessment and Technology: Mapping Future Directions in the Early Childhood Classroom," *Frontiers in Education*, vol. 4, 2019.
- [3] "FLIPCurric," March 2022. [Online]. Available: <http://flipcurric.edu.au/content>.
- [4] A.-M. Brady, "Assessment of learning with multiple-choice questions," *Nurse Education in Practice*, vol. 5, p. 238–242, 2005.
- [5] Y. Kats, "Learning Management Systems and Instructional Design: Best Practices in Online Education," 2013.
- [6] M. Douglas, J. Wilson and S. Ennis, "Multiple-choice question tests: A convenient, flexible and effective learning tool? A case study," *Innovations in Education and Teaching International*, vol. 49, p. 111–121, 2012.
- [7] D. R. Sadler, "Formative assessment and the design of instructional systems," *Instructional Science*, vol. 18, p. 119–144, 1989.
- [8] M. Heritage, "Assessment for learning as support for student self-regulation," *The Australian Educational Researcher*, vol. 45, February 2018.
- [9] D. Hodges, C. Eames and R. K. Coll, "Theoretical Perspectives on Assessment in Cooperative Education Placements," *Asia-Pacific Journal of Cooperative Education*, vol. 15, p. 189–207, 2014.
- [10] M. Escudier, T. Newton, M. Cox, P. Reynolds and E. Odell, "University students' attainment and perceptions of computer delivered assessment; a comparison between computer-based and traditional tests in a 'high-stakes' examination," *Journal of Computer Assisted Learning*, vol. 27, p. 440–447, 2011.



- [11] B. Foorman, M. York, K. Santi and D. Francis, "Contextual effects on predicting risk for reading difficulties in first and second grade," *Reading and Writing*, vol. 21, p. 371–394, 2008.
- [12] J. J. Montroy, T. A. Zucker, M. M. Assel, S. H. Landry, J. L. Anthony, J. M. Williams, H.-Y. Hsu, A. Crawford, U. Y. Johnson, M. S. Carlo and H. B. Taylor, "The Texas Kindergarten Entry Assessment: Development, Psychometrics, and Scale-Up of a Comprehensive Screener," *Early Education and Development*, vol. 31, p. 701–738, 2020.
- [13] V. E. Woloshyn, M. Bajovic and M. M. Worden, "Promoting Student-Centered Learning Using iPads in a Grade 1 Classroom: Using the Digital Didactic Framework to Deconstruct Instruction," *Computers in the Schools*, vol. 34, p. 152–167, 2017.
- [14] I. Arnold, "Cheating at online formative tests: Does it pay off?," *The Internet and Higher Education*, vol. 29, 2016.
- [15] "Online Classes See Cheating Go High-Tech," 2012. [Online]. Available: <https://www.chronicle.com/article/online-classes-see-cheating-go-high-tech/>.
- [16] O. R. Harmon, J. Lambrinos and J. Buffolino, "Assessment Design and Cheating Risk in Online Instruction," *Online Journal of Distance Learning Administration*, vol. 13, 2010.
- [17] T. B. Crews and D. F. Curtis, "Online Course Evaluations: Faculty Perspective and Strategies for Improved Response Rates," *Assessment & Evaluation in Higher Education*, vol. 36, p. 865–878, 2011.
- [18] U. S. A. Osuji, "The Use of e-Assessments in the Nigerian Higher Education System," *Turkish Online Journal of Distance Education*, vol. 13, p. 140–152, 2012.
- [19] J. Donovan, C. Mader and J. Shinsky, "Online vs. traditional course evaluation formats: Student perceptions," *Journal of Interactive Online Learning*, vol. 6, p. 158–180, 2007.
- [20] U. K. University of Southampton, N. Alruwais, G. Wills and M. Wald, "Advantages and Challenges of Using e-Assessment," *International Journal of Information and Education Technology*, vol. 8, p. 34–37, 2018.
- [21] J. Ridgway, S. Mccusker and D. Pead, "Literature review of e-assessment," 2004.
- [22] R. Ellaway and K. Masters, "e-Learning in medical education Guide 32 Part 2: Technology, management and design," *Medical Teacher*, vol. 30, no. 5, pp. 474–489, 2008.