The use of Canvas quizzes as low-stakes testing to facilitate language learning

Ruben Manuel Oronich

rmanuelo@uoc.edu

18/06/23

Tutor/Supervisor: Jennifer Rose Ament

Final Master's Thesis

Year 2022-2023, semester 2

Table of Contents

UOC Faculty
of Arts and
Humanities

Abstract

Quizzes are normally regarded as a tool to assess how much new content a student has attained after an arbitrary period of time. A quiz may take place at the end of a module, halfway through the quarter, etc. However, quizzes can also be used as low-stakes testing, a technique that tests learners periodically without awarding any grade. As a result, students are permitted to make mistakes and in turn they learn from them. If low-stakes testing is implemented on weekly basis, the effect of retrieval practice emerges. This process forces students to summon the information periodically and eventually the student learns the content of the quiz as a result of completing it. The objective of this FTM is therefore to evaluate the implementation of low-stakes testing by the means of online quizzes to improve the learning of Catalan as a second language at the University of California, Los Angeles (UCLA). Given that UCLA has recently adopted Canvas as the learning management system (LMS), the application of technology proposed in this project will be developed by using Canvas quizzes. To measure the effect of online quizzes on the learning of Catalan, the grades obtained by the current cohort of students exposed to this intervention will be compared to the grades of last year's cohort, acting as a control group. On top of that, the instructor will also conduct informal interviews with the students to obtain the students' point of view. The results of this implementation are expected to lead to successful learning and better grades at the end of the quarter. As a collateral benefit, students are also expected to notice an improvement in their learning habits.

Keywords: online quiz, low-stakes testing, retrieval practice, L2 Catalan

Resumen

Las pruebas normalmente se consideran una herramienta que evalúa cuánto contenido nuevo ha aprendido el alumnado después de un periodo de tiempo arbitrario, ya sea al final de un módulo, durante el trimestre, o en otro momento. Sin embargo, las pruebas también se pueden usar como evaluación de bajo riesgo, una técnica que evalúa a los estudiantes de forma periódica, pero sin recibir ninguna nota a cambio. Como consecuencia, el alumnado puede cometer errores y aprender de ellos. Si la evaluación de bajo riesgo se implementa semanalmente, emerge el efecto de la práctica de recuperación. Este proceso fuerza al alumnado a recuperar la información

periódicamente que termina aprendiendo el contenido de la prueba solo por el hecho de completarla. El objetivo de este TFM es evaluar la implementación de la evaluación de bajo riesgo en forma de pruebas en línea para mejorar el aprendizaje de catalán como segunda lengua en la Universidad de California, Los Ángeles (UCLA). Dado que UCLA acaba de adoptar Canvas como el sistema de gestión del aprendizaje, la herramienta de pruebas de Canvas se usará para la aplicación de tecnología propuesta en este proyecto. Para medir el efecto de las pruebas en línea en el aprendizaje de catalán, se compararán las notas obtenidas por el grupo de alumnos actual que ha completado estas pruebas con las del grupo del año pasado, que actúa como grupo de control. Además, el profesor o la profesora llevará a cabo entrevistas informales para obtener las opiniones del alumnado. Se espera que los resultados de la implementación conduzcan a un aprendizaje exitoso y mejores notas al final del trimestre. Como beneficio colateral, se espera que los estudiantes observen una mejora en sus hábitos de estudio.

Palabras clave: pruebas en línea, evaluación de bajo riesgo, práctica de recuperación, catalán como L2

UOC Faculty of Arts and Humanities
Universitat Oberta de Catalunya

1. Introduction

The pedagogical application of technology presented in this FTM seeks to analyze the use of online quizzes as a means of language learning and by providing an example of an implementation. As opposed to a more traditional approach where quizzes test language attainment at the end of a specific academic period (after certain number of hours, halfway through the course, at the end of the course, etc.), this project seeks to develop weekly online quizzes in Canvas to facilitate the process of learning Catalan as a second language in a higher education institution in the United States.

The use of online quizzes to generate learning can be explained with two concepts from the fields of second language acquisition and teaching: low-stakes testing and retrieval practice. Whilst the first term refers to a teaching-learning technique that involves quizzing students periodically without impacting the final grade, the second one relates to the cognitive process of summoning new information from the brain and its effects on boosting learning.

With this theoretical framework in mind, this proposal provides a detailed explanation of how to create online quizzes on Canvas and helps identifying the correct quiz option among those offered by Canvas and its configuration. In relation to this, the proposal looks into empirical evidence about quiz-related features to determine the most adequate number and type of questions, the relevance and provision of feedback, and test frequency, among other factors. In other words, past research is considered to create a Canvas quiz which will be used to *increase* learning instead of to *assess* learning.

The justification of this FMP is twofold. However, let me provide you with a little bit of context before going into detail. I started working as a Spanish and Catalan language lecturer at the University of California, Los Angeles (UCLA) in September 2021, where I have been teaching several language courses every quarter since then. Although the Spanish courses have been more diverse, the Catalan language units have been constant. Consequently, this research is motivated by two contextual factors I have encountered whilst teaching Catalan at this institution.

The first motive to conduct this study results from the combination of the arrival of Gen Z people, the first digital-native generation, to the university, and the adoption of Canvas as the default learning management system (LMS) at the university. This implementation has resulted in students being able not only to complete most of their learning in an

online format but also to interact with their peers, to contact their teachers, to submit papers, and to complete homework and examinations without leaving their working space. Technology has been part of these students' everyday life since birth, so it is crucial for institutions and instructors to accommodate these new students in the teaching-learning environment.

The second reason relates to the original design of the Catalan courses. When I first arrived, these units already had an online component in Canvas that students completed as part of their homework. These activities mainly consisted of a combination of language input presented in different formats (video, audio, and text) and a quiz to test its comprehension. The nature of these quizzes was to assess the language learned, exemplifying the most traditional and widespread use of them. Although this is not a bad characteristic per se, the quizzes were not designed to facilitate language learning, and its questions and format did not consider the indications and recommendations from the literature in the field.

2. Conceptual framework and objectives

The following section delves into the concepts of low-stakes testing and retrieval practice to frame this project within the field of second language acquisition and teaching. Then, a detailed explanation of the process to create a Canvas quiz is provided alongside a set of evidence that justify the format of the quiz.

It is worth noting that the empirical evidence presented below is mainly based on studies conducted outside the scope of SLA given the difficulty to find investigations that analyze the effect of low-stakes testing on language learning. For instance, a search on the UOC library webpage using the filter of *title contains* "low-stakes testing" and "language learning" provided no results. Another search with the same terms but widening the filter to *any field contains* leads to 120 results. Nonetheless, the studies presented do not focus on the effect of implementing low-stakes testing to learn a language but on other aspects related to the use of low-stakes testing such as students' anxiety, test validity and reliability, and other testing conditions.

2.1. Low-stakes testing and retrieval practice

According to the Center for Excellence in Learning and Teaching from the State University of New York (n.d.), low-stakes testing involves frequent assessment with

minimal effect on students' grades. The key element of this testing methodology is that students are allowed to try, make mistakes, and learn from them.

In their meta-analysis on the relationship between frequent low-stakes testing and class performance, Sotola and Crede (2021) highlight the popularity of this technique as a teaching tool among instructors of any field. Furthermore, frequent low-stakes testing is also well-received among students, and its main power relies on the fact that students need to periodically summon the information learned, a process that generates learning by itself. Picardo (2017) identifies that frequent low-stakes testing can have a positive effect on students' learning.

Batsell et al. (2017) analyzed the effect of low-stakes testing in the form of periodic quizzes among two groups of college psychology students. The control group, N=33, was assigned a daily reading and three exams over a semester. The intervention group, N=31, had to complete a daily quiz on top of the daily reading and the three exams. The authors concluded that low-stakes testing improved students' performance in comparison to the control group. The students who took the quizzes also reported having improved study habits for the exam by increasing their reading time, reducing cramming, and redistributing their study time. On a similar note, Kenney and Bailey (2021) analyzed the effect of testing content covered in previous lectures at the beginning of each class among 47 college students enrolled in a cognitive psychology course. The authors observed that quizzing had a positive impact on the performance in the final exam.

The impact of using low-stakes testing as a means to learn new informartion is tied to the concept of retrieval practice. In his summary about cognitive psychology research applied to education, Roediger III (2013) lists retrieval practice as one powerful strategy for learning, which emerged from an opposition to rote learning, a common teaching strategy at the beginning of the 1900's. According to the author, by practicing the retrieval of information, students will keep this information where it is accessible and therefore this information will be ready to be used.

Research conducted in the field shows that retrieval practice goes beyond the act of learning new content as it can also be a powerful memory enhancer (Roediger III & Butler, 2011). According to the authors, the process of searching for information as a result of the retrieval practice creates new associations between the target information

UOC Faculty of Arts and Humanities

and the concepts related to it (Sotola & Crede, 2021). That is, when the students are required to retrieve the information, they create a semantic network that increases the paths through which they are capable of accessing the information. These new retrieval routes facilitate access to this information in the future (Roediger III & Butler, 2011).

To test the effect of retrieval practice, Karpicke and Roediger III (2008) conducted a study on learning foreign language vocabulary word pairs. A group of college students were given a list of 40 pairs of Swahili-English words that they had to learn. The authors found out that those students who were repeatedly tested on the whole list of words experienced an improved long-term retention of words than those students who were only asked to study the pairs of words. The second group of students was not capable of retrieving the meaning of as many words as the intervention group. With similar results, Carpenter et al. (2008) observed that testing produced better results than just restudying 60 Swahili-English word pairs among 44 subjects drawn from their laboratory pool. The authors also concluded that tests maximize retention of information over time.

2.2. Online quizzes as a means to boost learning

Gen Z people, the first digital-native generation, constitute the majority of university students at the moment. This generation alongside the consequences of the implementation of online teaching during the COVID19 pandemic has played a significant role in the increase of use of online activities, assessment, materials, etc., in university courses. Persada et al. (2019) claim that these students have reshaped the teaching-learning environment by filling university rooms with computers, tablets, and electronic pencils. The use of this technology shapes how these students learn (Shatto & Erwin, 2016) and therefore the use of online quizzes, as opposed to a paper-version of them, should be seen as a natural next step.

Butler (2017) states that tests not only assess learning but also cause learning. As mentioned before, the process of retrieving information favors learning, and that is the process students undergo when asked to complete a quiz. Thanks to the quizzes, Smith and Weinstein (n.d.) claim that students are capable of identifying what they already know, and what they do not know yet. In other words, quizzing students increases their awareness of the learning process. In that regard Sotola and Crede (2021) highlight the importance of completing quizzes because it allows students to focus on new material once they can confirm that they know the covered content.

2.2.1. The impact of online quizzes

Ragupathi (2020) lists eleven advantages of using online quizzes. Among these, identifying areas of misunderstanding and misconceptions, providing detailed and immediate feedback, and providing repeated opportunities for learners to practice are some of the aspects students can benefit from by being quizzed online. Another set of positive elements identified by the author refers to the actual creation of the test and tackles aspects such as the ease of storing and reusability of quizzes and questions as well as the opportunity to shuffle and randomize questions.

The ability to provide feedback is also listed by Butler et al. (2008) as a benefit. However, the authors also put forth the ability to provide multiple attempts, the help students receive to learn the new content and the increase in students' engagement as some of the main advantages of online quizzes. In relation to students' engagement, Cann (2016) points out that in higher education faculty often complain about students' lack of motivation and implication with lessons. Thus, online quizzes might offer a good solution to that.

To exemplify how quizzes help students to learn, Johnson and Kiviniemi (2009) and Roediger (2013) note that students benefit from frequent practice that is distributed over time to prevent studying by cramming. In their study with 159 undergraduate students in an introductory psychology course, Johnson and Kiviniemi (2009) concluded that periodic testing impacted the students' final grade since they observed a positive correlation between the number of reading quizzes completed with the performance in class and the final course grade. Kibble (2007) provides evidence of the opposite effect in his study with medical physiology students. The author observed a significantly worse performance in the final examination from those students who did not complete the online quizzes. On a similar note, Sotola and Crede (2021) reached a similar conclusion by linking the completion of quizzes with higher academic performance. Furthermore, the authors claim that quizzing appears to be particularly useful for low-performing students.

Regarding favorable learning opportunities, Gamage et al. (2019) link the instant feedback provided on online quizzes to an improvement of understanding after conducting a research with undergraduate students of an engineering course. According to the authors, instant feedback is also valuable for the instructors since it provides them with a snapshot of the students' understanding of the topic.

In his study about the effects of testing and repeated studying on knowledge transfer with 48 undergraduate psychology students, Butler (2010) concludes that retrieval practice during testing promoted transfer. That is, the author unveiled that frequent testing is more effective than repeated studying, a conclusion already presented by Roediger III and Karpicke (2006). On a similar note, Roediger III and Butler (2011) claim that frequent testing has proven to benefit long-term retention.

2.2.2. Creating online quizzes

Ragupathi (2020) delves into six types of questions that may appear in an online quiz in relation to Bloom's taxonomy of education objectives. Bloom (1956) provides a scheme of six levels of learning that students need to navigate through, and Holtzman (2008) labels these six levels as follows: knowledge, comprehension, application, analysis, synthesis, and evaluation.

The six types of questions introduced by Ragupathi (2020) are the following: multiple choice questions, multiple response questions, fill-in-the-blanks, true-false questions, matching questions, and essay questions. It is worth mentioning that five out of these six types of questions are automatically checked; only the essay question needs to be graded manually. The types of questions, as well as their relation to the six levels of Bloom's taxonomy, are explained in Table 1.

Type of question	Description	Bloom's level of learning
Multiple choice question	Students need to select one correct answer from a list of possible options	Knowledge, comprehension, application, analysis
Multiple response question	Students need to select all the correct answers from a list of possible options	Knowledge, comprehension, application, analysis
Fill-in-the-blanks	Students are asked to add the answer, typically a one-word answer.	Knowledge, comprehension, application
True-False question	Students are asked to decide whether a statement is true	Knowledge, comprehension
Matching question	Students are asked to identify two related items	Knowledge, comprehension
Essay/short answer question	Students are asked to produce an answer which can range from a paragraph to several pages	Knowledge, comprehension,

		application, analysis, synthesis, evaluation

Table 1. Type of questions presented in Ragupathi (2020)

A more detailed look at multiple-choice tests is provided by Butler (2017). The author claims that this type of test is very popular because it is objective, easy to grade, and can cover a lot of content. As a supporter of multiple-choice testing, the author presents the three following recommendations as to how these tests should be created. First, the test should be challenging, but not too difficult. Allowing students to succeed prevents students from experiencing a negative effect from testing, such as learning the incorrect answer or being discouraged. But the author also notes that the test should not be too easy to complete because it would compromise the student retrieval process and sabotage their learning. Second, Butler claims that all three or four answers should be plausible. Although there is no agreement about the exact number of answers, too many would override the learning. And third, the author suggests not using the "none of the above" nor "all of the above" as an answer since these do not help students discriminate what is right or wrong. On the use of multiple-choice questions, Gamage et al. (2019) point out that setting up these types of questions is a cumbersome and time-consuming process that instructors need to be aware of.

Conversely, Sotola and Crede (2021) advocate for adding open questions so that students are asked to generate an answer as opposed to students answering a multiple-choice question. The authors explain that the cognitive processes involved in open and multiple-choice questions differ. In line with that, Roediger III and Butler (2011) suggest that production questions, rather than recognition questions, require more effort and therefore are better.

In relation to the number of questions that a quiz should have, ten seems to be the general rule. However, researchers also agree on the creation of a greater bank of questions. For instance, Cook and Banon (2017) conducted a study to analyze the use of online weekly quizzes based on prescribed readings among 458 students of a geography subject. Each quiz presented 10 questions to the students, but instructors were asked to create a total of forty questions for each of the mandatory readings.

Following the tenets of low-stakes testing, giving students multiple attempts to complete the test seems a right decision. There is some consistency among researchers that

students should have three or more attempts to complete an online quiz and that only the highest grade should be recorded. Nonetheless, in their study with 297 undergraduate students of a general education course called Theories of personality, DeSouza and Fleming (2003) gave students up to five attempts to complete each quiz although no reason is provided to explain their decision.

There is no agreement regarding the inclusion of feedback in low-stakes testing. Whereas Sotola and Crede (2021) claim that immediate feedback increases quiz effectiveness, Roediger III and Butler (2010) assure that the retrieval practice that occurs during testing is useful even without feedback. However, the authors later explain that showing the correct answer to the students boosts learning because it gives the opportunity for students to correct errors while maintaining the right responses. When it comes to multiple-choice and true/false questions, providing feedback to students is even more relevant since they have been exposed to incorrect information in the answers. In this particular case, the inclusion of feedback would prevent students from creating connections between the wrong answer and the prompt. In their review of laboratory studies, Roediger III and Karpicke (2006) highlight that feedback provided on quizzes allows students to learn from it, especially when it is elaborate and meaningful. Cann (2016) tested the effect of quizzing 315 first-year biological science students one week before each summative test. Among all the benefits listed, the author points out that he was able to include links to online resources and videos thanks to the online nature of the feedback and the formative nature of the quizzes.

Setting a time limit within which students need to complete the quiz is another aspect that researchers touch upon. Although the quiz length is intrinsically tight to the number and difficulty of questions and therefore it is difficult to find common ground, investigators acknowledge the necessity to have a limit. Hillman (2012) created a multiple-choice quiz for each of the twelve assigned readings for the undergraduate course Introduction to Abnormal Psychology that had 200 students. The author, who was also the instructor of the course, decided to set a 15-minute limit to complete the quiz. According to her, having a time limit forced students to read the materials in advance which helped them to focus on the actual test.

Test frequency and timing is another aspect worth addressing. Roediger III and Butler (2011) claim that when tests are conducted immediately after study, despite students producing very few errors, the retrieval effect is very poor. To explain this effect, the

authors suggest that immediate testing does not require too much effort for the student to access the information since the new material might not be stored in the long-term memory yet. Sotola and Crede (2021) claim that students should be tested in every class. However, most of the studies that tackle these issues are based on investigations that use quizzes as a preparatory tool, for example, students need to complete them before the next lecture. For example, Salas-Morera et al. (2012) analyzed the effect of quizzes to help students learn during a general education course for engineering students. Given that these quizzes were based on preparatory readings, students completed them on a weekly basis before each lecture.

Lastly, there is a controversy as to whether awarding a grade for the completion of online quizzes. Despite one of the tenets of low-stakes testing is that there should be no impact on the grade, the general ground is for quizzes to represent a small percentage of the final grade. Sotola and Crede (2021) underscore that quizzes may be more effective if they contribute towards the final grade of the unit so as students would then have an extrinsic motivator to complete them. For the studies already mentioned, the online quizzes on Johnson and Kiviniemi (2009) study accounted for only a 6% of the final grade whereas for Cook and Banon (2017) online quizzes represented the 20% of the final grade.

2.3.   Objectives and research question

The literature review presented indicates that the implementation of low-stakes testing facilitates learning, and it demonstrates that testing students periodically is more beneficial than asking them to study, or restudy, the course content. Alongside the empirical results, this methodology is also supported by the benefits of retrieval practice, a cognitive process that generates learning by asking students to summon information.

Given the current learning-teaching environment, the use of online testing is deemed to be the best tool to implement low-stakes testing. This tool offers many options, from presenting students with different types of questions to providing feedback or setting a time limit to complete a quiz. The fact that all learning management systems already allow instructors to create quizzes makes online quizzes the most convenient and cheapest tool to use.

As a result of the aforementioned, the aim of this pedagogical application of technology is to develop weekly online quizzes on Canvas that bolster the learning of Catalan as a L2 at UCLA. Therefore, this FMP seeks to answer the following question:

- How does the implementation of online quizzes as low-stakes testing facilitate the learning of Catalan as a second language in a higher education institution in the United States?

3. Methodology

3.1. Context and participants

The application of technology proposed in this FMP would take place in the module of Catalan Language and Culture 2 hosted by the Department of Spanish and Portuguese at the University of California, Los Angeles. If successful, the proposal will be later exported to the first Catalan language course.

UCLA follows the quarter system, and each of them lasts ten weeks. Since September 2021, Catalan 1 has been offered four times (Fall 21, Winter 22, Fall 22, and Winter 23) and there have been 4, 6, 7 and 10 students respectively. Catalan 2 has only been offered in two quarters (Spring 22, and Spring 23), with 2 and 3 students each time. The current cohort of Catalan 2 students consists of 2 males, aged 22, and 1 female, aged 23. They are all American nationals and speak English and Spanish. Two of them are considered Spanish heritage speakers and the other one learned the language in school. For two of them, Political Science is their major, whereas Neuroscience is for the other.

Lessons are scheduled twice a week (Monday and Wednesday) from 10am to 12pm, and the Catalan modules are not linked to a specific academic year. Therefore, students from any level can take them. Nonetheless, there is a prerequisite to confirm that there is a homogeneous level among the student cohort in Catalan 2: having passed Catalan Language and Culture 1 or having received an approval from the instructor after an interview. Learners enrolled in Catalan Language and Culture 1 aim to achieve an A1+ level at the end of quarter, and an A2+/B1 level is expected for Catalan 2 learners.

As a result of a departmental decision, these courses have two particularities worth mentioning. First, they are hybrid modules. Students from the campus in Los Angeles attend in-person, whereas students from any of the other ten campuses of the University of California complete the whole course synchronously via Zoom. For the last two years,

there have been students from the campus in Santa Barbara and Irvine, but this quarter all three students attend the lessons in-person. And second, both Catalan language and culture modules follow a flipped classroom approach, for which learners need to complete a set of online activities prior to each lesson.

3.2. Using Canvas to create online quizzes

After the decision from the University of California, Los Angeles to adopt Canvas as the university-wide learning management system, the implementation of low-stakes testing using online quizzes will be implemented in this LMS. To exemplify the application of technology presented in this FTM, one quiz has been created.
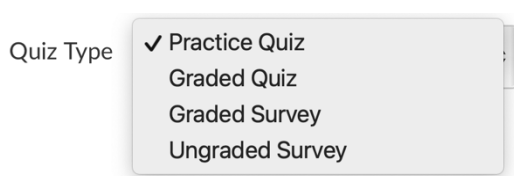
3.2.1. Types of quizzes and settings

As shown in Figure 1, Canvas offers four different types of quizzes divided into two main groups: quizzes and surveys.

Quiz Type
✓ Practice Quiz
Graded Quiz
Graded Survey
Ungraded Survey

Figure 1. Types of quizzes

According to the Canvas Instructor's guide, a *practice quiz* does not award any grade to the students and therefore it does not appear on the gradebook section of the Canvas page. However, students can see the points obtained based on their answers. This type of quiz is thought to help students monitor their learning. On the contrary, a *graded quiz* rewards students with points based on their responses, and Canvas automatically creates a column in the gradebook for each of this type of quiz.

As for the surveys, *graded surveys* are graded on completion but not on students' performance. That is, it does not matter how good or bad students complete the survey, just completing the survey awards points, and the results also appear on the gradebook. Finally, *ungraded surveys* are planned to obtain opinions or other information from students without providing a grade.

For this FTM, the option of *practice quiz* will be used. The motivation of this choice relies on its alignment with the aforementioned characteristic of low-stakes testing: minimal effect on students' grades. The practice quiz will allow students to complete it, see how well they perform but it will not have an impact on their grade.

There are six options and restrictions for *practice quizzes* in Canvas. If selected, the option *Shuffle answers* randomizes the answers for multiple choice questions. *Time limit*

allows the instructor to set a maximum number of minutes for the students to complete the quiz. Adding a specific number of attempts and deciding which score is recorded (highest, last, or average) is activated by selecting *Allow multiple attempts*. If -- appears on the box for allowed attempts, it means that students can take the quiz as many times as they want. Students will be able to see their responses and receive automated feedback for correct or incorrect answers as well as know which questions they got wrong if the option *Let Students See Their Quiz Responses* is selected. For this option, students may be restricted to seeing their answers *Only Once After Each Attempt.* If the student is allowed multiple attempts, quiz responses may be made available after *Only Once After Each Attempt* or *Only After Their Last Attempt*. The opportunity to *Let Students See The Correct Answers* is also available. Correct answers might be displayed after each attempt or during a selected period of time. The last option available is to *Show one question at a time.* If selected, students will be presented with one question at a time, and the instructor may also decide to *Lock questions after answering* so that the student cannot go back to it. Although these are not relevant for this research, Canvas also allows the addition of two restrictions. First, quizzes may *Require an access code*, and second it may *Filter the IP address.* Finally, the instructor can also decide whether to assign the quiz to a whole group or to a selected group of students. Moreover, the due date as well as the availability of the quiz can also be set.

Figure 2. Quiz settings

Following the indications provided in the literature review, the settings of the example quiz created for this proposal are shown in Figure 2. The quiz has a time limit of 20 minutes, and students have three attempts to complete it. Given that the results of the quizzes do not impact the grade, the highest score will be recorded by default. Students will be able to see their answers as well as the correct answers after each attempt so that they can learn from the feedback provided.

UOc Faculty of Arts and Humanities
Universitat Oberta de Catalunya

3.2.2. Creating the questions

The quiz created to exemplify the technological application proposed contains twelve questions, which is very close to the 10-question quiz that tends to dominate the research. Canvas offers the following types of questions: multiple choice, true/false, fill in the blank, fill in multiple blanks, multiple answers, multiple dropdowns, matching, numerical answer, formula question, essay question, file upload question, and text (no question). The quiz uses 5 different types of questions: multiple answers, multiple dropdowns, fill in multiple blanks, multiple choice, and file upload questions, four of which are graded automatically. This will help decrease the time the instructor spends on correction. To create the questions, the indications provided by Ragupathi (2020) have been considered.

Figure 3. Multiple answer question

Figure 3 shows a multiple-answer question. In this question, students need to recall the regular forms of the past participle and are asked to select all the correct answers. This type of question fulfils the Knowledge objective proposed by Bloom's taxonomy presented in Ragupathi (2020).

A multiple dropdowns question is shown in Figure 4. This type of question resembles a

Figure 4. Multiple answer question

multiple-choice question since students are asked to select the one correct answer from different answers. However, on Canvas this type of question allows the presentation of several prompts in the same question. This question also fulfills the Knowledge objective from Bloom's taxonomy (Ragupathi, 2020) since it asks

students to identify the answer.

An example of a fill in multiple blanks is shown in Figure 5. In this case, this question asks students to produce the answer. For this specific question, students need to understand the subject of the sentence and conjugate the verb accordingly. Therefore, it relates to the Comprehension category of Bloom's Taxonomy (Ragupathi, 2020).

Figure 5. Fill in multiple blanks question

The next type of question used in the quiz is multiple choice. Figure 6 shows two questions that assess the understanding of the use of the tense, where students must relate sentences written in *passat perfet* and *passat perifràstic* with the correct adverbial of time. According to Bloom's Taxonomy, it falls into the Analysis category (Ragupathi, 2020). As instructed by Butler (2017), the questions are challenging enough for students but not too difficult to discourage them. Moreover, the options "none of the above" and "all of the above" are not used.

Figure 6. Multiple choice questions

Finally, the last type of question used is a file upload question. Students are asked to produce a short voice note (see Figure 7) in which they need to explain what they did today using the new tense. According to Bloom's Taxonomy, this is an Application question (Ragupathi, 2020). This last question alongside the fill-in-the-blanks questions is presented to fulfill Roediger III and Butler (2011) and Sotola and Crede (2021) claims.
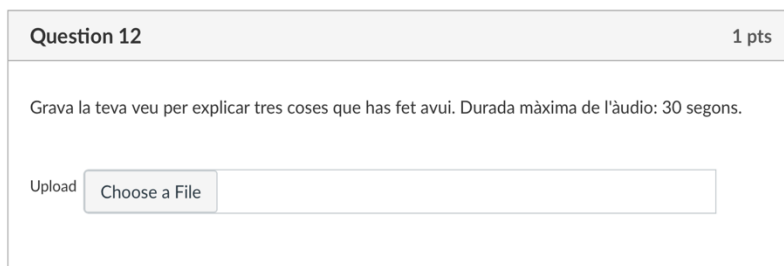
Image 7. File upload question

Open questions were students need to generate an answer require more effort and different cognitive processes are involved. As a consequence, the authors claim that these are better.

With regards to the inclusion of feedback, Image 8 shows an example. Following Roediger III and Karpicke (2006), feedback is provided so that students can learn from it. For the correct answer, positive feedback is given so that the student settles their learning and reinforces their understanding. For the negative answer, the comments give a hint of what the students should be aware of (talking about a time that has finished, *ahir* (*yesterday*) and la *setmana passada* (*last week*), as opposed to a time that is still related to the present, *aquest matí* (*this morning*)) and also provide a link to another resource that student may want to check before completing the next attempt.
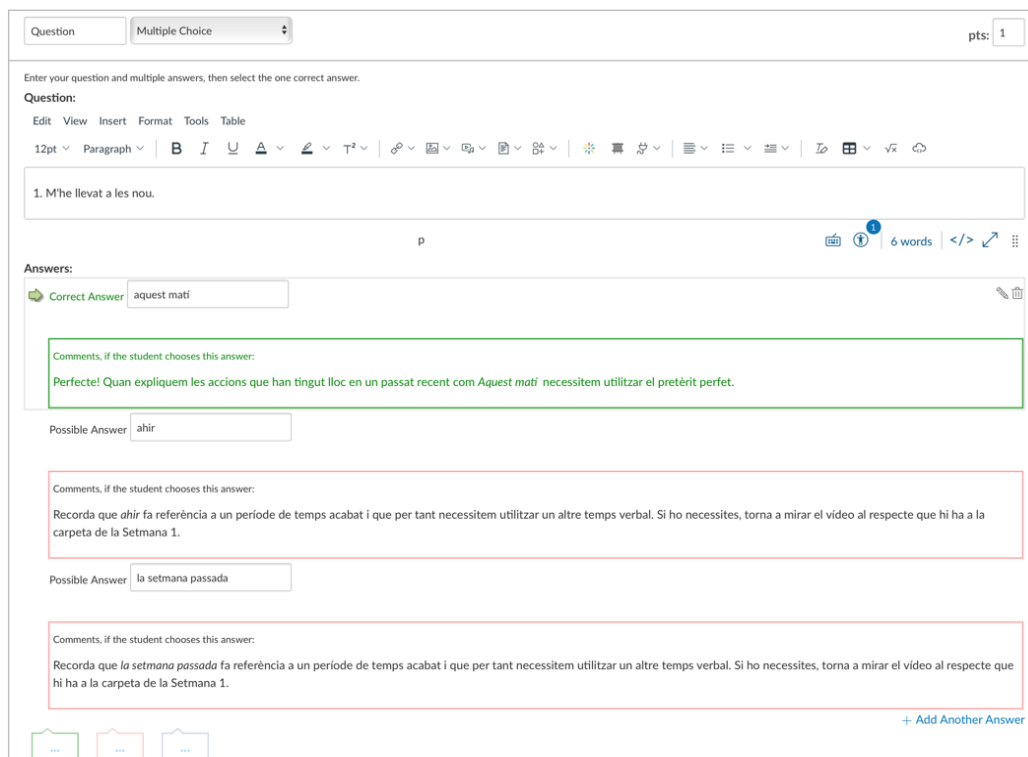

Image 8. Feedback

3.3. In-class application

This FMP tests the use of weekly quizzes on 3 learners of Catalan in a level 2 course. In addition to the regular course material a weekly quiz has been designed and administered over the course of 9 weeks to see if their completion has any effect on the learner's linguistic knowledge. More specifically, these quizzes test the knowledge of the Catalan grammar.
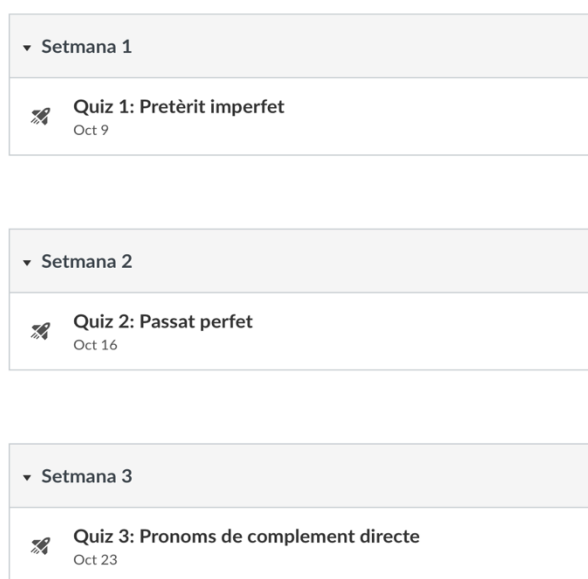

Figure 9. Schedule of quizzes

Following Roediger III and Butler (2011) findings, quizzes will not be completed just after learning the new language since the retrieval effect will be very low. Therefore, the quiz based on the content already taught will be made available to students on Friday morning, two days after the last class of the week, and it will have to be submitted before the following class on Monday. Figure 9 shows the distribution of the tests during the first three weeks. For students to complete the quizzes, they just need to log onto the Canvas page and access the quiz by clicking on its name.

3.4. Measuring the effects of the implementation

The outcome of the use of online quizzes to facilitate language learning will be measured in two different ways. First, an analysis of the quiz effect by looking at the students' final grades can be conducted. Ideally, the results obtained from the treatment group should be compared to the results obtained from a control group, a group of students who would follow the exact same curriculum except for completing the quizzes. However, given the low number of students in the Catalan 2 cohort, the analysis of the effects of quizzes will be done by comparing the current students' grade with that of last year's students.

Second, the instructor may conduct informal chats with all the students to gather their opinions about the quizzes' usefulness. The instructor will have the responsibility to determine the quizzes' impact by asking students about its difficulty, timing, possible improvements, etc.

3.5. Predicted outcomes of the implementation

Based on the research mentioned in the literature review, this section aims to answer the research question presented in this FTM:

- How does the implementation of online quizzes as low-stakes testing facilitate the learning of Catalan as a second language in a higher education institution in the United States?

I predict that the implementation of weekly quizzes would have a positive impact on students' learning and on final exams results in line with Kenney and Bailey (2021). In addition, should a student not complete the weekly quizzes but complete the rest of the course, I expect a poorer performance in the final exam compared to their peers as Kibble (2007) noted.

From the informal interviews, I anticipate a good reception of this technique among students as Sotola and Crede (2021) suggest. In line with Batsell et al. (2017), I also project that students would observe an improvement of their learning habits. It is expected that students use the quizzes to identify what they know and what they need to know ahead of their final evaluation.

From observing how the students behave in class, a backlash on using online quizzes is not expected. As a matter of a fact, the three students who participated in this study use their computer as their sole learning tool, from completing the in-class activities to submitting other homework.

4. Discussion and conclusion

This proposal sought to unveil the effect of implementing online quizzes as low-stakes testing to facilitate the learning of Catalan as a foreign language. With the empirical research presented in this paper as well as with the expected results discussed above, it seems likely that the implementation of low-stakes testing would generate a successful result in the students' process of learning the language.

Thanks to the information provided in the literature review, the quizzes and their settings should be adjusted accordingly. For instances, quizzes would be completed on a weekly basis following the tenets of low-stakes testing. Moreover, the time between the last lesson of the week and the weekly quiz would be set at two days to enhance the retrieval process of information. Likewise, the quiz itself would be created following the general

agreements detailed in the literature review. For example, the quiz should have multiple attempts, it should include feedback for both right and wrong answers and it should have a variety of question types, which would force students to use different cognitive processes to complete it.

From what has been said, it is easy to conclude that this implementation would likely be successful. To list only some of the benefits and strengths, bringing this proposal to the real teaching environment would pose no difficulties. Once the main syllabus is created, the instructors would only need to create one weekly test on the same LMS that they are using. The fact that it is not necessary to find an external software makes the instructor's job very easy and there is no need for the department or university to spend extra money on this new software. Quizzes are also easy to store and edit if necessary. Automatic grading is another strength of this implementation because it reduces correction time during the quarter.

Conversely, one clear limitation involves the initial time that instructors need to occupy to create the quizzes. Without help, this can be a cumbersome process for the main instructor. However, as mentioned above, once the quizzes are created it is very easy to edit them if needed.

One aspect that should be considered in the future is the impact of these quizzes on the final grade. Although this proposal underscores the idea that low-stakes testing should not affect the final grade, research shows that students' willingness and motivation to complete these quizzes increases if the results account for a percentage of the final grade. Even if it is minimal, this can be a powerful motivator for students. Therefore, I believe these quizzes should award students a small percentage of the final grade. Finally, it would be important to conduct this study with larger cohorts and with different languages, levels and contexts to determine the real impact of using online quizzes as low-stakes testing.

UOC Faculty of Arts and Humanities
Universitat Oberta de Catalunya

References

Batsell Jr, W. R., Perry, J. L., Hanley, E., & Hostetter, A. B. (2017). Ecological validity of the testing effect: The use of daily quizzes in introductory psychology. *Teaching of Psychology, 44*(1), 18-23. https://doi.org/10.1177/0098628316677492

Bloom, B. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain.* David McKay Company.

Butler, A. (2017, October 10). *Multiple-choice Testing: Are the Best Practices for Assessment Also Good for Learning?* The Learning Scientists. https://www.learningscientists.org/blog/2017/10/10-1

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1118-1133. https://doi.org/10.1037/a0019902

Butler, M., Pyzdrowski, L., Goodykoontz, A., & Walker, V. (2008). The effects of feedback on online quizzes. *International Journal for Technology in Mathematics Education*, *15*(4).

Cann, A. J. (2016). Increasing student engagement with practical classes through online pre-lab quizzes. *Journal of Biological Education*, *50*(1), 101-112. https://doi.org/10.1080/00219266.2014.986182

Carpenter, S.K., Pashler, H., Wixted, J.T., & Vul. E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition, 36*, 438-448. https://doi.org/10.3758/MC.36.2.438

Center for Excellence in Learning and Teaching (n.d.). *Low-stakes testing*. https://www.oswego.edu/celt/low-stakes-testing

Cook, B. R., & Babon, A. (2017). Active learning through online quizzes: better learning and less (busy) work. *Journal of Geography in Higher Education*, *41*(1), 24-38. https://doi.org/10.1080/03098265.2016.1185772

DeSouza, E., & Fleming, M. (2003). A comparison of in-class and online quizzes on student exam performance. *Journal of Computing in Higher Education, 14*, 121-134. https://doi.org/10.1007/BF02940941

Gamage, S. H., Ayres, J. R., Behrend, M. B., & Smith, E. J. (2019). Optimising Moodle quizzes for online assessments. *International journal of STEM education*, *6*(1), 1-14. https://doi.org/10.1186/s40594-019-0181-4

Holtzman, M. (2008). Demystifying application-based multiple-choice questions. *College Teaching*, *56*(2), 114-120. https://www.jstor.org/stable/20695189

Johnson, B. C., & Kiviniemi, M. T. (2009). The effect of online chapter quizzes on exam performance in an undergraduate social psychology course. *Teaching of Psychology*, *36*(1), 33-37. https://10.1080/00986280802528972

Karpicke, J. D., & Roediger, H. L., III. (2008). The critical importance of retrieval for learning. *Science, 15*, 966-968. https://doi.org/10.1126/science.1152408

UOC Faculty of Arts and Humanities

Kenney, K. L., & Bailey, H. (2021). Low-stakes quizzes improve learning and reduce overconfidence in college students. *Journal of the Scholarship of Teaching and Learning, 21*(2). https://doi.org/10.14434/josotl.v21i2.28650

Kibble, J. (2007). Use of unsupervised online quizzes as formative assessment in a medical physiology course: effects of incentives on student participation and performance. *Advances in Physiology Education*, *31*(3), 253-260. https://doi.org/10.1152/advan.00027.2007

Persada, S. F., Miraja, B. A., & Nadlifatin, R. (2019). Understanding the Generation Z Behavior on D-Learning: A Unified Theory of Acceptance and Use of Technology (UTAUT) Approach. *International Journal of Emerging Technologies in Learning*, *14*(5). http://dx.doi.org/10.3991/ijet.v14i05.9993

Picardo, J. (2017, July 9). *How to do IT: Using digital technology to support effective assessment and feedback*. https://my.chartered.college/impact_article/how-to-do-it-using-digital-technology-to-support-effective-assessment-and-feedback/

Ragupathi, K. (2020). *Designing effective online assessments. Resource guide.* National University of Singapore. https://www.nus.edu.sg/cdtl/docs/default-source/professional-development-docs/resources/designing-online-assessments.pdf

Roediger III, H L., & Butler, A. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 5*(1), 20-27. https://doi.org/10.1016/j.tics.2010.09.003

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological science*, *17*(3), 249-255. https://www.jstor.org/stable/40064526

Roediger III, Henry L. (2013). Applying Cognitive Psychology to Education: Translational Educational Science. *Psychological Science in the Public Interest, 14*(1), 1-3. https://psycnet.apa.org/doi/10.1177/1529100612454415

Salas-Morera, L., Arauzo-Azofra, A., & García-Hernández, L. (2012). Analysis of online quizzes as a teaching and assessment tool. *Journal of Technology and Science Education, 2*(1), 39-45. https://doi.org//10.3926/jotse.30

Shatto, B., & Erwin, K. (2016). Moving on from millennials: Preparing for generation Z. *The Journal of Continuing Education in Nursing*, *47*(6), 253-254. https://doi.org/10.3928/00220124-20160518-05

Smith, M., & Weinstein, Y. (n.d.). *The Many Benefits of Retrieval Practice*. https://www.learningscientists.org/blog/2016/5/5-1

Sotola, L.K., & Crede, M. (2021). Regarding Class Quizzes: a Meta-analytic Synthesis of Studies on the Relationship Between Frequent Low-Stakes Testing and Class Performance. *Educational Psychology Review, 33*, 407-426. https://doi.org/10.1007/s10648-020-09563-9

UOC Faculty of Arts and Humanities
Universitat Oberta de Catalunya

Appendix I – Canvas quiz

Question 1   3 pts

Quin dels següents infinitius té un participi regular?

☐ anar - anat

☐ fer - fet

☐ perdre - perdut

☐ entendre - entés

☐ dormir - dormit

Question 2   5 pts

Relaciona cada infinitiu amb la forma del participi correcta.

entendre --> [Select] ⬍

obrir --> [Select] ⬍

fer --> [Select] ⬍

tenir --> [Select] ⬍

voler --> [Select] ⬍

Question 3   2 pts

Quins són els participis correctes?

☐ venir - vinut

☐ beure - begut

☐ escriure - escrigut

☐ creure - cregut

Question 4   9 pts

Escriu els participis dels infinitius següents:

anar -->

fer -->

aconseguir -->

entendre -->

dormir -->

dir -->

sentir -->

rebre -->

conèixer -->

Question 5
5 pts

Completa les oraciones amb la forma del passat perfet del verb entre parèntesis.

1. La Marta [   ] (escriure) una carta a la Maria a mà.

2. La Maria [   ] (parlar) tres hores per telèfon amb el Miquel.

3. La Carla [   ] (entendre) tots els exercicis de gramàtica.

4. L'Oriol [   ] (viure) tres anys a l'Argentina.

5. L'Arnau [   ] (obrir) la porta d'una esbategada.

Question 6
2 pts

Marca les frase en passat perfet.

☐ La Núria estudiava tota la tarda.

☐ El Toni ha arribat a casa.

☐ L'Ignasi va fer tots els deures.

☐ El Rafa ha obert la porta.

Relaciona cada frase amb el connector adequat:

Question 7   1 pts

1. M'he llevat a les nou.

○ aquest matí

○ ahir

○ la setmana passada

Question 8   1 pts

2. Em vaig llevar a les nou.

○ aquest matí

○ ahir

○ demà

Question 9   1 pts

3. He arribat a Maó.

○ fa dos dies

○ aquesta tarda

○ abans-d'ahir

Question 10
1 pts

4. Vaig arribar a les Illes Balears

○ fa dos dies

○ aquesta tarda

○ aquest cap de setmana

Question 11
2 pts

Transforma les frases següents al passat perfet.

1. La Maria i el Joan van tenir un fill. -->

2. L'Olga i tu vau arribar tard. -->

Question 12
1 pts

Grava la teva veu per explicar tres coses que has fet avui. Durada màxima de l'àudio: 30 segons.

Upload   Choose a File

UOc Faculty of Arts and Humanities