applied sciences

MDPI

Article

A Case Study to Explore a UDL Evaluation Framework Based on MOOCs

Francisco Iniesto [1,*], Covadonga Rodrigo [2] and Garron Hillaire [3]

[1] Institute of Educational Technology, The Open University, Milton Keynes MK7 6AA, UK
[2] School of Computer Science, Universidad Nacional de Educación a Distancia (UNED), 28015 Madrid, Spain
[3] Advanced Education Research and Development Fund, Oakland, CA 94612, USA
* Correspondence: francisco.iniesto@open.ac.uk

Abstract: In this paper, we focus on 23 undergraduate students' application of a universal design for learning (UDL) evaluation framework for assessing a massive open online course (MOOC) in the context of a usability and accessibility university course. Using a mixed-methods approach, we first report the extent to which untrained raters agree when evaluating their course with the framework and then examine their feedback on using UDL for assessment purposes. Our results indicate user feedback provides great value for both the future development of accessible MOOCs and identifies opportunities to improve the evaluation framework. For that purpose, we suggest an iterative process comprised of refining the framework while working with students and which could help students to internalise UDL principles and guidelines to become expert learners and evaluators. The complexities and redundancies that surfaced in our research, as reported in this paper, illustrate that there is variability in the perception of both the course design and the interpretation of the framework. Results indicate that UDL cannot be applied as a list of simple checkpoints, but also provide insights into aspects of the framework that can be improved to make the framework itself more accessible to students.

Keywords: massive open online course; universal design for learning; course design; accessibility; evaluation framework; mixed methods

Citation: Iniesto, F.; Rodrigo, C.; Hillaire, G. A Case Study to Explore a UDL Evaluation Framework Based on MOOCs. *Appl. Sci.* **2023**, *13*, 476. https://doi.org/10.3390/app13010476

Academic Editor: Habib Hamam

Received: 24 October 2022
Revised: 2 December 2022
Accepted: 27 December 2022
Published: 29 December 2022

1. Introduction

The objectives of UNESCO's sustainable development goal 4 (SDG4) is to ensure inclusive, equitable, and quality education and to promote lifelong learning opportunities for all. This has proven to be a challenge in online learning, and in particular in open educational resources (OERs) and massive open online courses (MOOCs) [1]. While MOOCs have marked a significant shift in online learning, they offer a huge range of open-access courses to the public; most people who enrol in MOOCs already have a graduate-level education and many of the enrolled students do not complete the course [2]. The fact that MOOCs are available to a global audience is a positive aspect, but they must be accessible to everyone, regardless of their needs [3].

Through the research presented in this paper, we aim to contribute to make MOOCs beneficial to all students by focusing on the learning design and examining if it is accessible. For that purpose, we understand that user feedback is important for the future development of accessible MOOCs. Therefore, we use YourMOOC4all (YourMOOC4all, http://yourmooc4all.lsi.uned.es/ accessed on 1 September 2022), a recommender system which allows any student to freely evaluate a MOOC to see if it meets the principles of universal design for learning (UDL) [4]. The use of UDL in education offers both students and educators benefits by removing barriers to learning through giving all students the same opportunity to achieve their learning goals [5]. The application of UDL in primary, secondary, and tertiary contexts is widespread and growing. It already plays a significant role in university curricula [6] and is used in educational international initiatives [7,8].

The objective of this research was (1) to evaluate the accuracy of the UDL evaluation framework by untrained raters and (2) their perceptions of the usefulness of UDL as an evaluation framework to identify accessibility barriers. With this intention, we have collected feedback from 23 third-year computer science (CS) undergraduates taking part in a usability and accessibility university course.

2. Background

2.1. UDL as an Evaluation Framework for MOOCs

The UDL framework is comprised of three design principles that contain nine guidelines and 31 checkpoints (see the Appendix A for the structure of the framework). The principles specify the overall goal, while the checkpoints supply design suggestions considering universal design in learning contexts. Students are differently motivated to learn and perceive the educational content; some students are more interested in the process of learning and others are more interested in the results of learning, while others work differently during learning [9]. Therefore, the UDL approach is to present the information in ways that are easy to understand for students, rather than forcing them to adapt to the information [10,11].

MOOCs offer a way for more people to get involved in learning. For example, recent research shows that there are benefits for students regardless of their background when taking MOOCs [12]. These courses are relatively affordable, making them a great option for students for continuing professional development (CPD) [13] and facilitating equity, diversity, and inclusion (EDI) values in education [14]. MOOCs are designed to be student-centred, and so to benefit from them, students must be prepared to work by investing time in their learning (Handoko et al., 2019). It is relevant to reflect on the learning design of MOOCs and their technical accessibility, and to understand how these elements are affecting participation and completion rates [15].

In terms of accessibility evaluation, some accessibility guidelines for online courses, such as the web content accessibility guidelines (WCAG) (Web content accessibility guidelines (WCAG) 2.1., https://www.w3.org/TR/WCAG21/ accessed on 1 September 2022), can be difficult to assess because of the limitations of current accessibility standards, for example, regarding the evaluation of learning disabilities [15,16]. Unfortunately, there are few references in the literature that discuss students' expectations concerning accessibility and what they would like to improve in MOOCs [17]. We have found that there is a critical aspect of inclusive design, which is often ignored in MOOCs, that is needed for detailed accessibility information to ensure that students with accessibility needs can fully access the online learning platform and its educational resources [4].

The UDL framework is designed to produce educational content that is based on its principles, rather than being used to evaluate educational content [18]. According to recent research, using UDL to classify and address accessibility barriers in online learning is a sound approach [19]. UDL is aligned with the pedagogical perspective of MOOCs, where students are expected to be self-directed in their learning, wherein the objective of UDL is to help novice learners become expert learners by mastering the learning process. [20,21]. As stated, the UDL framework promotes the building up of expert learners. According to Iniesto and Hillaire [22], using the UDL framework for MOOCs assessment helps students understand technology accessibility and how to learn effectively. Participants can benefit from evaluating MOOCs by becoming expert learners and evaluators.

2.2. YourMOOC4all

YourMOOC4all is a joint research project between The National Distance Education University (UNED) and The Open University (OUUK) which contains MOOCs in Spanish from Coursera (Coursera, https://www.coursera.org/ accessed on 1 September 2022), UNED Abierta (UNED Abierta, https://iedra.uned.es/ accessed on 1 September 2022), and MiriadaX (MiriadaX, https://miriadax.net/cursos accessed on 1 September 2022). Similarly to other MOOC search engines, such as Class Central (Class Central https://www.

class-central.com/ accessed on 1 September 2022) or CourseTalk (CourseTalk https://www.coursetalk.com/ accessed on 1 September 2022), it allows students to provide feedback on the MOOCs they are taking part in and to be recommended other courses based on their CPD interests. YourMOOC4all offers a valuable feature for MOOC students: the opportunity to review the MOOCs' learning experience, through ratings and free text comments. Its design is developed on the premise that students' experiences on learning platforms provide useful feedback to feed other students' interests and accessibility needs (see Figure 1).

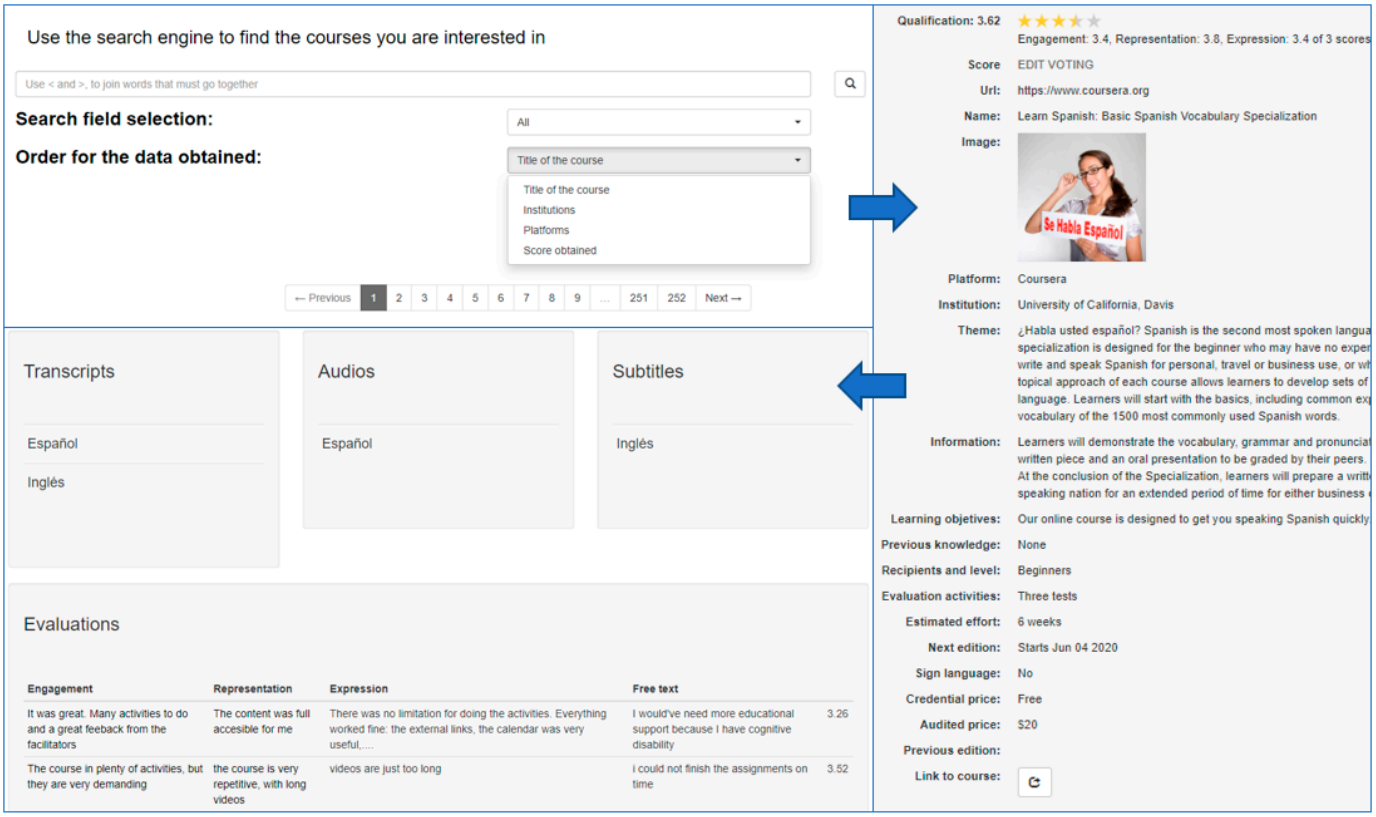

Figure 1. YourMOOC4all key functionalities.

For the course reviews, UDL is used. For that purpose, an evaluation checklist was created following UDL guidelines [9]. The evaluation checklist created by the authors includes 31 questions directly related to UDL checkpoints. Students can use a Likert scale to rate any of the optional indicators using 0 to 5. The indicators within the checklist offer some helpful insights when it comes to answering each question (see Figure 2). In the evaluation process, students can provide qualitative feedback which enriches the quality of the feedback, proportionate information to other students, and generates data to help identify accessibility barriers to MOOC providers. The complete set of questions is included in the Appendix A.

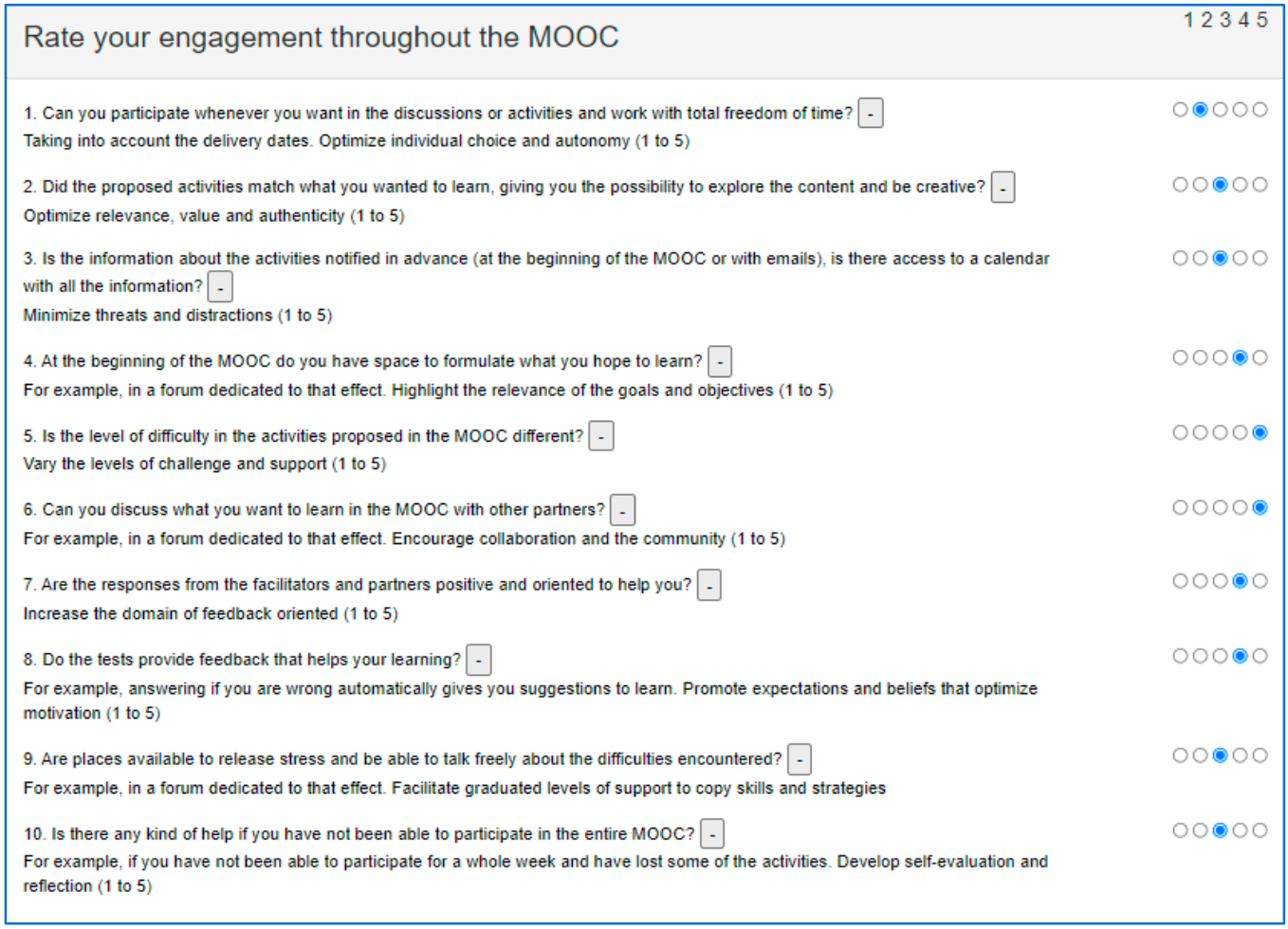

Figure 2. Provide multiple means of engagement checklists.

3. Methodology

3.1. University Course and Sample

The context of this study was the *"Usability and Accessibility"* (Usabilidad y accesibilidad) course, which is part of the computer engineering degree at UNED. Third-year CS undergraduates are introduced to the guidelines for designing accessible graphical user interfaces, developing accessible webpages, and implementing the use of automatic and manual tools and methodologies for assessing web accessibility (i.e., the use of The World Wide Web Consortium (W3C) standards (W3C, https://www.w3.org/ accessed on 1 September 2022). The course has two assignments to address continuous assessment. The second one is an in-depth study of WCAG guidelines and accessibility evaluation where undergraduates are asked to assess the accessibility of the MOOC "*Accessible digital materials*". This MOOC is designed to develop students' skills for the development of accessible learning resources and the identification of accessibility barriers [23]. This blended pedagogical approach allows students to assess the accessibility of the MOOC while they participate in an external educational resource which covers similar topics to the university course [24].

During the academic course 2018–2019, an optional exercise was included in the second assignment, where students used YouMOOC4all. In the assignment, students first had to evaluate the accessibility of the requested MOOC through WCAG guidelines and then come up with the evaluation of the MOOC using the UDL framework. The experience included a sample of 33 students enrolled in the course (86% male and 93% Spanish), from which 23 students answered the optional exercise (70%).

3.2. Objectives and Research Questions (RQs)

As stated above, MOOCs, if accessible, have a great potential for developing CPD and EDI values in education and the use of the UDL framework promotes the building up of expert learners and evaluators. With this intention, we have collected feedback from third-year CS undergraduates with experience in the evaluation of web accessibility (i.e., WCAG) but not in UDL. For that purpose, undergraduates use YourMOOC4all to assess the same MOOC using the proposed UDL framework.

The two objectives of this research conducted with undergraduates were (1) to evaluate how accurate and easy it is to understand and use the UDL evaluation framework by untrained raters (i.e., non-expert evaluators):

1. RQ1. To what extent did untrained raters agree when using the UDL evaluation framework?

In addition, (2) their perceptions of the usefulness to assess accessibility barriers using the UDL evaluation framework included in YourMOOC4all:

2. RQ2. What are the perceptions of UDL as an evaluation framework for untrained raters?

3.3. Methods

As reported by Myers and Powers [25], a mixed-methods approach allows for a deeper and broader perspective of the phenomena researched, formulates the problem statement more clearly, and finds the best way to approach it, both theoretically and practically, by producing varied data through a multiplicity of observations. The methodology is designed to gather differentiated but rich data considering the limited sample. Therefore, two sources of data were designed for this research:

1. The Likert and open questions existing in YourMOOC4all to assess a MOOC using the UDL framework (quantitative and qualitative).
2. A new set of open questions included in the exercise script (qualitative).

Table 1 summarises the two tasks delivered to students, task 1, to answer RQ1, included the first source of data. While task 2 incorporated the second source of data to support RQ2.

Table 1. YourMOOC4all exercise summary.

YourMOOC4all Exercise	
(Task 1) Process (accompanied by screenshots in the script). (RQ1)	• Step 1. Search for "*Accessible digital materials*" in YourMOOC4all search engine • Step 2. Select the course in the search engine to be evaluated • Step 3: 1. Evaluate UDL in its three principles using the checkpoints (Likert—1 strongly disagree, 2 disagree, 3 neutral, 4 agree, 5 strongly agree). 2. Enter your evaluation in the open-ended questions (open question). • Step 4. Save the evaluation in YourMOOC4all.
(Task 2) Questions to answer in the script. (RQ2)	1. Reflect on the advantages and disadvantages of using UDL (open question). 2. Find out which aspects of WCAG 2.1 does not evaluate UDL and vice versa (open question). 3. Comment on which checkpoints you have detected as difficult to evaluate (open question). 4. Comment on what checkpoints seem redundant (open question).

For the analysis of the quantitative data, inter-rater reliability was tested using Fleiss's kappa [26]. Fleiss' kappa is a measure to assess the reliability of ratings between a fixed number of people when assigning ratings to several categories. The measure calculates how much different ratings are classified in a way that is not due to chance. In this case, the selected Fleiss' kappa is fixed-marginal multi-rater because students were assigned a set number of cases to each category (i.e., the Likert scale).

While for the open questions, the method of thematic analysis was selected for analysis [27]. Thematic analysis is a way of looking at data that involves identifying

patterns in meaning across them, considering the authors' experiences when looking at data to create a more complete and accurate understanding of the subject matter. The thematic analysis process involved question-responses read by the authors and coded. Then the authors reviewed potential themes using references and frequencies. Finally, the themes were compared with the original data to see if they were appropriate for interpretation. Names from students have been made anonymous using ST (from "*student*") and a number.

4. Results

4.1. RQ1. To What Extent Did Untrained Raters Agree When Using the UDL Evaluation Framework?

The results of the interaction of undergraduates with YourMOOC4all have been divided first by checkpoints, then by guidelines and principles, in each of the following figures. The mean, standard deviation (SD), and Kappa (K) are shown on the diverging stacked bar charts. K Interpretation is 0.0–0.20 slight agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 substantial agreement; and 0.81–1.0 almost perfect agreement [26]. Two Kappa values have been calculated, K1 includes the five Likert values, while K2 is reduced to three options (disagreement, neutral, and agreement). Fair agreement values are presented with a * while moderate, substantial, and perfect agreements are shown with a + to facilitate the visibility of the results. Results are complemented by a sample of quotes from the open-ended questions during the evaluation using YourMOOC4all.

In the case of "*provide multiple means of engagement*" (Figure 3), in the MOOC, students identify they can participate in the discussions or activities and that the responses from the facilitators are positive and oriented to help (checkpoints 8.3 and 8.4), for example:

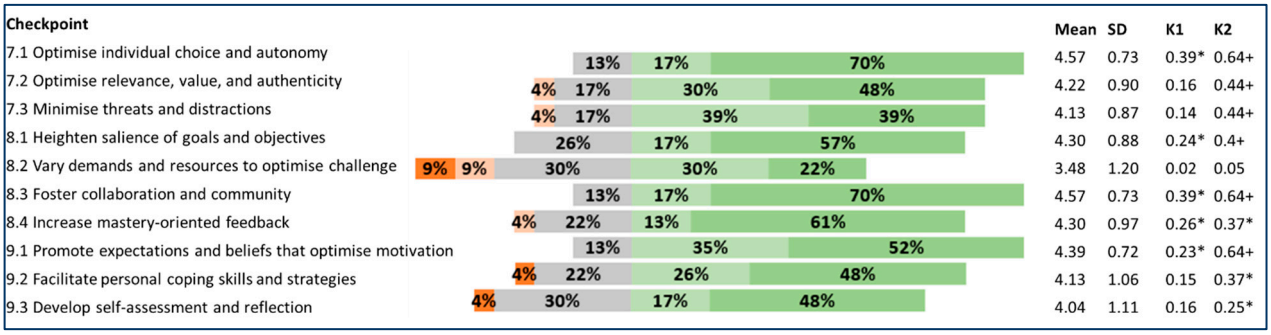

Figure 3. "Provide multiple means of engagement" evaluation results. Note: standard deviation (SD), Kappa for 5 values (K1); Kappa for 3 values (K2); fair agreement *; moderate, substantial, and perfect agreements +.

> *There is a forum where you can contact your classmates and thus release stress and continue learning thanks to their help. The tests contain great feedback on what was taught, but do not identify its level of difficulty. As a help, there is only one glossary, with certain terms and the forum for the "team" to answer your questions.* (ST8)

Students agree MOOC is designed to allow motivation and activities to match with the learning outcomes, with information for optimising individual choice and options for self-regulation (9.1 and 9.3):

> *The course is designed to effectively motivate the student. Its structure does not only seek purely theoretical content but plays with various options to achieve a key motivation so that students can develop their activities, ask their questions and progress in the content in an even fun way.* (ST7)

Concerning "*provide multiple means of representation*" (Figure 4) students are positive about videos containing captions and transcripts, the use of the language consistent terminology and having a logical sequential ordering of tasks (1.2 and 2.1):

I think that the representation of contents throughout the course is done in a good way, with the information provided in different formats and styles to allow everybody access to it. (ST20)

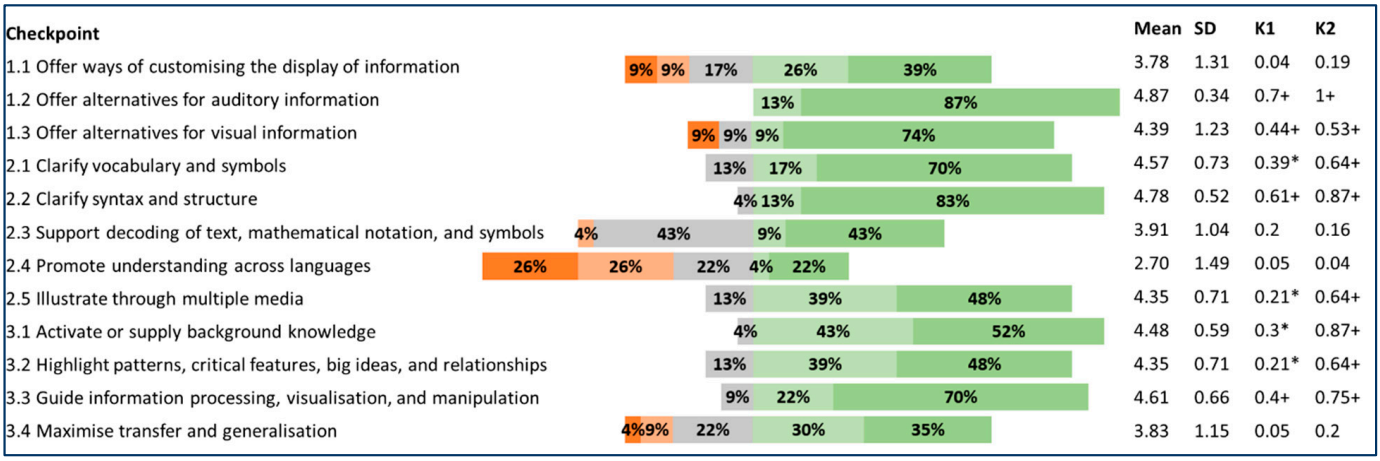

Figure 4. "Provide multiple means of representation" evaluation results. Note: standard deviation (SD), Kappa for 5 values (K1); Kappa for 3 values (K2); fair agreement *; moderate, substantial, and perfect agreements +.

While the MOOC and its platform are failing to adapt to the environment, modify the information and personalise the learning experience (1.1 and 3.4):

The content seems to me to be presented concisely. At all times you see the content index, which lets you know where you are going and not disconnect from the conceptual map of the course. The "weak" points of the MOOC are, for example, that it does not allow for modification of the visualisation of the content. (ST23)

Regarding the principle "*provide multiple means of action and expression*" (Figure 5, students are not confident with the use of social networks or external tools available (5.1):

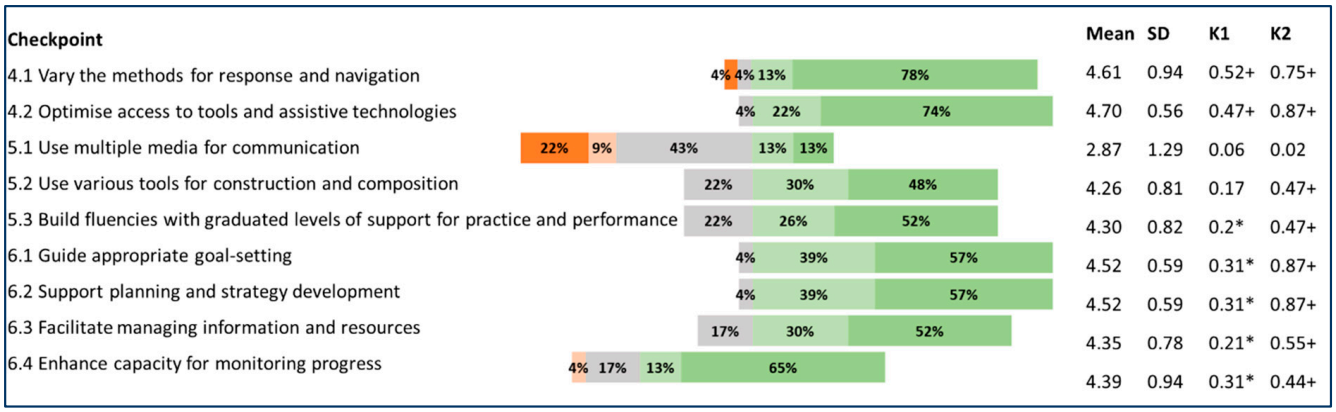

Figure 5. "Provide multiple means of action and expression" evaluation results. Note: standard deviation (SD), Kappa for 5 values (K1); Kappa for 3 values (K2); fair agreement *; moderate, substantial, and perfect agreements +.

According to the scope of the expression in the MOOC, the only point I find regarding the proposed evaluation is the non-existence of the possibility of communication through social networks (or at least I have not been able to find it anywhere in it). (ST27)

Students understand the MOOC is supporting the process of reflection, the availability of information, and the capacity for monitoring progress (6.4):

I found the course progress screen very interesting; it is very useful since it allows you to better control the time you have to finish it. (ST 6)

The results in terms of principles and guidelines have been described in detail above, but can also be analysed in aggregate form. The evaluation is generally positive and shows fair and moderate agreements (Figures 6 and 7), being the worst-rated guideline, and least agreed, "*expression and communication*" (5).

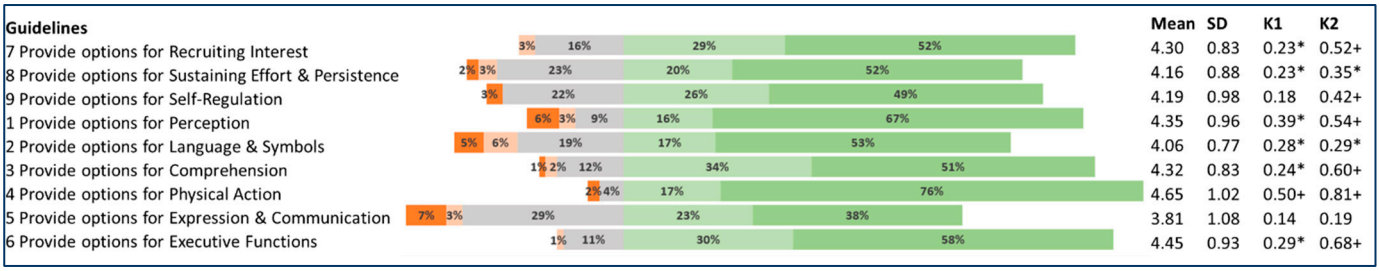

Figure 6. Guidelines evaluation results. Note: standard deviation (SD), Kappa for 5 values (K1); Kappa for 3 values (K2); fair agreement *; moderate, substantial, and perfect agreements +.

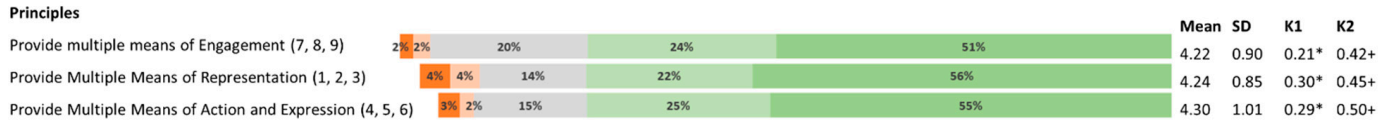

Figure 7. Principles evaluation results. Note: standard deviation (SD), Kappa for 5 values (K1); Kappa for 3 values (K2); fair agreement *; moderate, substantial, and perfect agreements +.

To answer RQ1, Fleiss' kappa values were computed for both K1 using the five Likert values in the questions, and K2 which reduced the evaluation to three options (disagreement, neutral, and agreement). For the 31 checkpoints using K1 scores there were: 11 slight, 17 fair, and 2 moderate agreements. In addition, the 31 checkpoints using K2 scores were: 6 slight, 5 fair, 17 moderate, and 2 substantial agreements. These results indicate that while the agreement for some items was achieved (i.e., 2.2. and 4.2), for other items, the responses among raters were variable (i.e., 1.1, 2.3, 2.4, 5.1, 5.3, and 8.2). The lower levels of agreement can be interpreted either as variable insights into limitations of course design or could be an indication of different interpretations of the evaluation tasks.

It is important to recall that UDL aims to design up front to consider the variability of students [5]. In that sense, in our research questions, the focus is on interpreting the results from the perspective that variable ratings represent the variability of students (in RQ2 we examine the potential for different interpretations of the evaluation framework). The notion of designing with consideration for human variability is that the design decisions that are necessary for some students are beneficial for all students. From this perspective, all areas where students disagree are potential opportunities for improvement in course design. The relationship between disagreement and agreement evaluations provides a potential prioritisation mechanism to address design concerns. Across all checkpoints, the results indicated there were 14 out of 31 checkpoints where at least one student disagreed, indicating the course did not implement the UDL checkpoint (see Figures 2–5). Of the 14 checkpoints with disagreement evaluations, 11 of those had slight agreement ratings using K1 scores (i.e., 1.1, 2.3., 2.4, 3.4, 5.1, 5.2, 7.2, 7.3, 8.2, 9.2, and 9.3) whereas 6 had slight agreement ratings using K2 scores (i.e., 1.1, 2.3, 2.4, 5.1, 5.3, and 8.2).

The strengths and limitations of prioritising course improvements using agreement statistics of course evaluations are bound to the frequency of disagreement [28]. Prioritising the six K2 slight agreement checkpoints would encompass all checkpoints with at least 10% of students with disagree evaluations. While it would help improve the overall evaluation for many students, it might not identify issues of critical importance that were identified by small numbers of students. It would be important to reconcile prioritisation by considering

which groups of students would benefit from the revisions. Minority groups of students may be also in the minority in terms of their UDL evaluations.

As can be seen in Table 2, with the strengths and limitations of prioritising course improvement using agreement statistics in mind, these results suggest focusing course improvement on six checkpoints where there was slight agreement using the K2 calculation.

Table 2. Checkpoints with a slight agreement.

Principles	Checkpoints
Provide multiple means of Engagement (7, 8, 9)	8.2 Vary demands and resources to optimise challenge
Provide Multiple Means of Representation (1, 2, 3)	1.1 Offer ways of customising the display of information 2.3 Support decoding of text, mathematical notation, and symbols 2.4 Promote understanding across languages 3.4 Maximise transfer and generalization
Provide Multiple Means of Action and Expression (4, 5, 6)	5.1 Use multiple media for communication

4.2. RQ2. What Are the Perceptions of UDL as an Evaluation Framework for Untrained Raters?

Table 3 details the thematic analysis including codes and quantification of the student's responses to the questions included in the second task of the exercise (question 1 (Q1) is divided between advantages and disadvantages).

Table 3. Codes derived from students' perceptions of UDL.

Question	Codes
Advantages and disadvantages (Q1)	• Advantages: Equity, Diversity, and Inclusion (5), Expectations and motivations (3), Learning design and assessment (3), Alternative formats (2), Language (1) • Disadvantages: Difficult to implement (2), Complex (2), Expensive (1), Unfamiliar (1)
Comparison (Q2)	Universal Design (4), Accessibility (3), Expectations and motivations (1), Usability (1), Personalisation (1)
Difficulty to evaluate (Q3)	Overlap between checkpoints (3), Communication (2), Learning design and assessment (2), Alternative formats (1), Personalisation (1), Self-regulation (1)
Redundancy (Q4)	Alternative formats (3), Communication (2), Learning design and assessment (2), Language (2) Personalisation (2), Time limit (1)

Advantages. We could anticipate that participants would see the value of designing up front for student variability. Furthermore, common beliefs about UDL were expected to appear in student responses. Those expectations were confirmed because the predominant categories of EDI, Expectations and Motivations, Learning Design, and Alternative Formats are features UDL implementation seeks to accomplish. An example of a response that illustrates the alignment of student-perceived advantages and the UDL framework is as follows:

> *UDL optimises learning so that in a group where we find students of different levels and abilities, we can teach everyone equally without excluding them. Facilitates access to study material, offering access in more than one format. In this way, it also promotes motivation among students and their participation.* (ST13)

Disadvantages. As we had no clear expectations of how untrained raters would interpret and use the checklist, the disadvantages help establish what work remains in operationalising UDL as an evaluation framework. Students identified the checklist as difficult to implement, complex, and time-consuming. Students also indicated a lack of familiarity with the UDL framework. An example quote that illustrates the challenges is as follows:

There may always be a student who cannot use the created product; therefore, it is necessary to design strategies and curricula that are inclusive for as many students as possible. Despite this, some students will need individualised support and attention. And despite everything, the main disadvantage that UDL brings is the large investment that must be made in educational centres and the little interest on the part of public and private institutions to carry it out. (ST9)

Comparison. Students were asked to compare the use of WCAG and UDL. While students in the sample are familiar with WCAG, UDL was new to them. Students understand WCAG as a set of guidelines for web accessibility but lack the pedagogical perspective included in UDL. Students have the perception that WCAG is included to some extent in some of the UDL guidelines, specifically when using the new version of WCAG (2.1) since the new criteria are oriented to accessibility on multiple devices. However, WCAG is designed to correct technical aspects, whereas UDL is for the design and evaluation of pedagogical aspects:

WCAG 2.1 are more oriented to the correction based on the staging of the content, and to the variety of tools and the good use of them, without presenting errors in their implementation, to facilitate user access. UDL is a methodology that values more conceptually the mechanisms that promote learning and make it more open to a greater number of people. (ST18)

Difficulty to evaluate. Students identified several checkpoints as difficult to evaluate (see Table 4), indicating the overlap between checkpoints during the evaluation.

Table 4. Checkpoints identified as difficult to evaluate by students.

Principles	Checkpoints
Provide multiple means of Engagement (7, 8, 9)	7.2 Optimise relevance, value, and authenticity 8.1 Heighten salience of goals and objectives 8.2 Vary demands and resources to optimise challenge 8.4 Increase mastery-oriented feedback 9.1 Promote expectations and beliefs that optimise motivation 9.2 Facilitate personal coping skills and strategies
Provide Multiple Means of Representation (1, 2, 3)	1.1 Offer ways of customising the display of information 2.3 Support decoding of text, mathematical notation, and symbols 3.4 Maximise transfer and generalisation
Provide Multiple Means of Action and Expression (4, 5, 6)	5.3 Build fluencies with graduated levels of support for practice and performance 6.2 Support planning and strategy development

Students report how difficult some checkpoints are formulated to evaluate without being strongly engaged with the MOOC considering aspects such as the learning design, assessment, or communication, which include checklists assessing the role of facilitators and interaction with other students and aspects related to learning outcomes and adaptation of the content:

The checkpoints where it is assessed whether the proposed activities agree with what it is desired to learn are difficult to assess since it depends on each of the students. It is the same case of the level of difficulty of the MOOC activities, the feedback in the tests and the existence of questions that help reflection. (ST14)

Redundancy. Regarding redundancy, students report that several checkpoints ask about similar concepts, in some cases, redundancy is within the principle such as in Groups 1 and 2 (see Table 5) for an evaluation of the use of language and monitoring progress. An example includes:

The checklists about discussing with students what you want to learn are redundant. In the case of the existence of a social network or external tool, the MOOC already has enough tools to be able to work with it. (ST7)

However, other identified redundancies exist across multiple principles, which make it more difficult to simplify the evaluation framework and show the possible overlaps in UDL within guidelines belonging to different principles. An example quote that shows redundancy across checklists and guidelines is:

The different questions about the language could be unified since they are redundant. The questions about which tools are used within the MOOC are also repetitive. Finally, a couple of times we are asked about the content, formats, and structures of the MOOC. (ST8)

Table 5. Checkpoints identified as redundant by students.

Group	Across Principles	Checkpoints	Description of Redundancy
Group 1	No	2.1 Clarify vocabulary and symbols 2.2 Clarify syntax and structure	The evaluation of the use of language regarding the existence of a glossary of terms (2.1) and maintaining the same terminology (2.2)
Group 2	No	6.2 Support planning and strategy development 6.4 Enhance capacity for monitoring progress	Monitoring progress and strategy development facilitating reflection (6.2) and progress (6.4), both in the quizzes
Group 3	Yes	1.2 Offer alternatives for auditory information 1.3 Offer alternatives for visual information 2.5 Illustrate through multiple media 5.1 Use multiple media for communication 5.2 Use various tools for construction and composition	Use of alternative formats. The inclusion of different questions formats such as captions and transcripts (1.2), audio descriptions (1.3), images, text, video, or graphics (2.5), the use of external tools (5.1), and external links (5.2) adds confusion
Group 4	Yes	4.1 Vary the methods for response and navigation 7.1 Optimise individual choice and autonomy	Time limit (4.1) is related to options for physical action, allowing extra time to achieve a task, (7.1) indicates options for recruiting interest allowing the needed time to participate in discussions
Group 5	Yes	5.3 Build fluencies with graduated levels of support for practice and performance 8.1 Heighten salience of goals and objectives 8.3 Foster collaboration and community 8.4 Increase mastery-oriented feedback	The interaction in the system and between peers for reflection. Including facilitators (5.3) space to formulate and share learning objectives (8.1), with other partners (8.3) and again facilitators (8.4)
Group 6	Yes	6.1 Guide appropriate goal setting 7.3 Minimise threats and distractions 9.2 Facilitate personal coping skills and strategies	Information about learning objectives, activities (6.1), the use of a calendar (7.3), and spaces to discuss difficulties encountered (9.2)

To answer RQ2, for the 31 checkpoints, 11 were identified as difficult to evaluate (see Table 4), and 18 were associated with a redundant group (see Table 5). At the intersection of difficult-to-evaluate and redundancy, there are 5 checkpoints (i.e., 5.3, 6.2, 8.1, 8.4, and 9.2). This suggests that for 20 of the 31 checkpoints, students did not find it difficult to evaluate. Students also did not see ambiguity for 13 of the 31 checkpoints. There is a distinction between how difficult a task is and how accurate a student is at the task. Just because something is hard does not necessarily mean that it was performed incorrectly.

Further insight is gained in that some of the checkpoints were both identified as difficult to evaluate and considered redundant with other checkpoints. This suggests that there is room to improve the language around the checkpoints for evaluation to

reduce the ambiguity for students. There must be certain redundancy and overlap within the framework too; the key characteristic is that many of the checkpoints reported as redundant belong to different guidelines and even principles. Perceptions of students are that UDL is useful and has benefits, but that using the checklist is not straightforward and training and experience for its application are needed. While there is some ambiguity and some areas that are difficult to evaluate the fact that students identify this as beneficial suggests this evaluation framework should be iterated on and improved to better support student evaluations.

5. Discussion

The evaluation related to RQ1 indicated there were six checkpoints with a slight agreement: 1.1, 2.3, 2.4, 3.4, 5.1, and 8.2 (see Figures 3–5) because participants provided a broad range of evaluation responses. To support all students, these checkpoints are a good focus for design revisions for the course. Further insight was gained around these checkpoints with results from RQ2. Students identified 11 checkpoints that were difficult to evaluate (see Table 4). The intersection of checkpoints with slight agreement and checkpoints that are difficult to evaluate were four of the six checkpoints (i.e., 1.1, 2.3, 3.4, and 8.2). This intersection suggests the range of evaluation scores may be due to the difficulty to evaluate the checkpoint for these four items. In contrast, checkpoints 2.4 and 5.1 had a slight agreement and were not identified as difficult to evaluate. This shows that the range of responses is more likely due to an accurate range of opinions about the course design. Therefore, the results indicate that the next steps in improving the course should focus on improving design decisions related to checkpoints 2.4 and 5.1.

Further insights arise from RQ2 related to the ambiguities the students identified in the evaluation framework. At the intersection of ambiguity, the slight agreement indicates that checkpoint 5.1 was considered ambiguous with checkpoints 1.2, 1.3, 2.5 and 5.2 (see Table 5). This would suggest that when considering design decisions to improve the course based on checkpoint 5.1, the course designers may gain more design insights by considering the related checkpoints. Table 6 summarises the checkpoints recommended for revision.

Table 6. Checkpoints suggested for revision.

Checkpoints
1.2 Offer alternatives for auditory information
1.3 Offer alternatives for visual information
2.4 Promote understanding across languages
2.5 Illustrate through multiple media
5.1 Use multiple media for communication
5.2 Use various tools for construction and composition

The main limitation of the proposed framework is that UDL is intended to be used in the design process while producing educational content [18]. The experiment has shown that it is challenging to be in the role of a student evaluating the course since every participant has a different individual perspective on aspects such as level of difficulty, reflection, and feedback. These aspects indicate the need to empower students for improving and refining the quality of the checkpoints included in YourMOOC4all [29]. That is aligned with the complexity and redundancy of the UDL evaluation framework as reported by the students, the number of indicators to evaluate in the framework is quite high (31), and students felt it was a time-consuming task.

The potential of using UDL for the evaluation of MOOCs has been previously reported [22]. The feedback provided in this study through ranked and open questions has proven useful to indicate how UDL used as an evaluation framework provides feedback for the inclusive design of online learning environments. Raters in this research knew about accessibility and specifically about WCAG evaluation but were untrained in evaluating with UDL. Some of the findings from this case study reveal common criticism made to universal design in general and UDL in particular: the lack of perception that some students may need

a user-centred approach [30] acknowledging not all are necessarily expert learners. In a MOOC environment, there exists a lack of support from the educational team, with only a few facilitators for a big ratio of students [31]. In that sense, UDL, if well-designed, can be a starting point to provide extra individual support.

6. Conclusions

As a limitation of this research, we understand that even the rich amount of data gathered from a sample of 23 students is not large enough to generalise the results. As well, other research methods and types of analysis for comparison could have been considered. Therefore, as discussed, future research should focus on removing redundancies and simplifying the evaluation questionnaire. Further studies should scale up the number of participants with varied backgrounds and interests. The inclusion of a control or comparison group made up of students who are not enrolled on the usability and accessibility course should be considered to compare the results. Finally, further research methods such as interviews and observations could be considered, as well as different types of analysis for the quantitative data to increase reliability.

This research has shown students have variable needs. Even with just 23 students, we have seen that variation. The goal of UDL is to design up front considering student variability [5]. This research has explored the intersection of MOOC design and student variability through the UDL expert evaluation framework. We have demonstrated a student-centred strategy to close the gap between design and evaluation by benefiting from the perceptions of CS undergraduate students who are not expert raters but have knowledge of accessibility. The process has shown that students have variable viewpoints on the checkpoints and have variable criticism of the course design which indicates that UDL cannot be applied as a list of effortless checkpoints.

Author Contributions: Conceptualization, F.I. and C.R.; Formal analysis, F.I. and G.H.; Investigation, C.R.; Methodology, F.I. and G.H.; Project administration, C.R.; Writing—original draft, F.I.; Writing—review & editing, C.R. and G.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by CLARA-HD grant number PID2020-116001RB-C32.

Institutional Review Board Statement: UNED Ethics committee approval 2-ETSIIF-TFG-2022.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: We are grateful for the time and dedication of the participants in this study. We want to thank Alejandro Rodríguez-Ascaso and Emilio Letón-Molina for the access to the MOOC Accessible digital materials, Shailey Minocha for proofreading the article and Pedro Manuel Muñoz for programming the prototype.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. UDL principles, guidelines, checkpoint items and checkpoint items adapted as questions.

Provide Multiple Means of Engagement	**Provide Multiple Means of Representation**	**Provide Multiple Means of Action and Expression**
Provide options for Recruiting Interest (7) • Optimise individual choice and autonomy (7.1) Can you participate whenever you want in the discussions or activities and work without time limits? • Optimise relevance, value, and authenticity (7.2) Did the proposed activities match what you wanted to learn, giving you the possibility to explore the content and be creative? • Minimise threats and distractions (7.3) Is the information about the activities notified in advance (at the beginning of the MOOC or with emails), is there access to a calendar with all the information?	Provide options for Perception (1) • Offer ways of customizing the display of information (1.1) Is it possible to adapt the environment to your needs, modifying the information that appears? • Offer alternatives for auditory information (1.2) Are there captions and transcripts available in the videos? • Offer alternatives for visual information (1.3) Are there audio descriptions available in the videos?	Provide options for Physical Action (4) • Vary the methods for response and navigation (4.1) Is there a time limit to perform the tests or activities when you start them? • Optimise access to tools and assistive technologies (4.2) Is it possible to move around the MOOC using only the keyboard or the mouse?
Provide options for Sustaining Effort & Persistence (8) • Heighten salience of goals and objectives (8.1) Do you have space to formulate what you are expecting to learn at the beginning of the MOOC? • Vary demands and resources to optimise challenge (8.2) Is the level of difficulty in the activities proposed in the MOOC differentiated? • Foster collaboration and community (8.3) Can you discuss what you want to learn in the MOOC with other partners? • Increase mastery-oriented feedback (8.4) Are the responses from the facilitators positive and oriented to help you?	Provide options for Language & Symbols (2) • Clarify vocabulary and symbols (2.1) Is the use of the language simple and understandable, also, is there a glossary of the terms used during the MOOC? • Clarify syntax and structure (2.2) Is the structure of the MOOC similar and maintains the same style, using the same terminology? • Support decoding of text, mathematical notation, and symbols (2.3) Are the mathematical terms clarified using a list of terms or a glossary? • Promote understanding across languages (2.4) Is the use of different languages supported? • Illustrate through multiple media (2.5) Are the most important concepts within the MOOC available in various formats such as images, text, video, or graphics?	Provide options for Expression & Communication (5) • Use multiple media for communication (5.1) Are there social networks or external tools available in the MOOC? • Use multiple tools for construction and composition (5.2) Are external links and complementary readings offered in the MOOC? • Build fluencies with graduated levels of support for practice and performance (5.3) Do the MOOC facilitators help in the process of communication and reflection?

Table A1. *Cont.*

Provide Multiple Means of Engagement	Provide Multiple Means of Representation	Provide Multiple Means of Action and Expression
	Provide options for Comprehension (3)	Provide options for Executive Functions (6)
	• Activate or supply background knowledge (3.1)	• Guide appropriate goal-setting (6.1)
Provide options for Self-Regulation (9)	Are the most important concepts in the MOOC explained at the beginning of it?	Is it clear at the beginning of each module what is to be learned and the calendar of activities?
• Promote expectations and beliefs that optimise motivation (9.1)	• Highlight patterns, critical features, big ideas, and relationships (3.2)	• Support planning and strategy development (6.2)
Do the tests provide feedback that helps your learning?	If there is a need for prior knowledge, is this indicated?	Are there quizzes during the MOOC to facilitate reflection on what has been learned?
• Facilitate personal coping skills and strategies (9.2)	• Guide information processing and visualisation (3.3)	• Facilitate managing information and resources (6.3)
Is there a space available to talk freely about the difficulties encountered?	Is the sequential ordering of tasks in the MOOC logical?	Are guides provided to assist in the learning process and the use of the platform?
• Develop self-assessment and reflection (9.3)	• Maximise transfer and generalisation (3.4)	• Enhance capacity for monitoring progress (6.4)
Is there any help in case you have not been able to participate in the whole MOOC?	Does the MOOC provide tools to personalise your experience and generalise learning?	Does the MOOC show the progress you have made?

References

1. Iniesto, F.; Tabuenca, B.; Rodrigo, C.; Tovar, E. Challenges to achieving a more inclusive and sustainable open education. *J. Interact. Media Educ.* **2021**. [CrossRef]
2. Deng, R.; Benckendorff, P.; Gannaway, D. Learner engagement in MOOCs: Scale development and validation. *Br. J. Educ. Technol.* **2020**, *1*, 245–262. [CrossRef]
3. Barrera, A.G.; Hernández, P.G.; López, C.M. La atención a la diversidad en los MOOCs: Una propuesta metodológica. *Educ. XX1* **2017**, *20*, 215–233.
4. Iniesto, F.; Rodrigo, C. YourMOOC4all: A recommender system for MOOCs based on collaborative filtering implementing UDL. In *European Conference on Technology Enhanced Learning*; Springer: Cham, Switzerland, 2019; pp. 746–750.
5. Meyer, A.; Rose, D.H.; Gordon, D.T. *Universal Design for Learning: Theory and Practice*; CAST Professional Publishing: Lynnfield, MO, USA, 2014. [CrossRef]
6. Gronseth, S.; Dalton, E.; Khanna, R.; Alvarez, B.; Iglesias, I.; Vergara, P.; Ingle, J.C.; Pacheco-Guffrey, H.; Bauder, D.; Cooper, K. Inclusive Instructional Design and UDL Around the World. In Proceedings of the Society for Information Technology & Teacher Education International Conference, Las Vegas, NV, USA, 18 March 2019; pp. 2357–2359.
7. Bracken, S.; Novak, K. (Eds.) *Transforming Higher Education Through Universal Design for Learning: An International Perspective*; Routledge: London UK, 2019.
8. Griful-Freixenet, J.; Struyven, K.; Verstichele, M.; Andries, C. Higher education students with disabilities speaking out: Perceived barriers and opportunities of the Universal Design for Learning framework. *Disabil. Soc.* **2017**, *32*, 1627–1649. [CrossRef]
9. CAST. Universal Design for Learning Guidelines Version 2.2. Wakefield, MA. 2018. Available online: http://udlguidelines.cast.org (accessed on 1 September 2022).
10. Cook, S.C.; Rao, K. Systematically applying UDL to effective practices for students with learning disabilities. *Learn. Disabil. Q.* **2018**, *41*, 179–191. [CrossRef]
11. Rose, D.H.; Meyer, A. The future is in the Margins: The Role of Technology and Disability in Educational Reform. In *A Practical Reader in Universal Design for Learning*; ERIC: Harvard Education Press: Cambridge, MA, USA, 2006; pp. 13–37.
12. Hung, C.Y.; Sun, J.C.Y.; Liu, J.Y. Effects of flipped classrooms integrated with MOOCs and game-based learning on the learning motivation and outcomes of students from different backgrounds. *Interact. Learn. Environ.* **2019**, *27*, 1028–1046. [CrossRef]
13. Naidu, S.; Karunanayaka, S.P.; Ariadurai, S.A.; Rajendra, J.C.N. Designing Continuing Professional Development MOOCs to promote the adoption of OER and OEP. *Open Prax.* **2018**, *10*, 179–190.
14. Lambert, S.R. Do MOOCs contribute to student equity and social inclusion? A systematic review 2014–2018. *Comput. Educ.* **2020**, *145*, 103693. [CrossRef]

15. Iniesto, F.; McAndrew, P.; Minocha, S.; Coughlan, T. Accessibility in MOOCs: The stakeholders' perspectives. In *Open World Learning: Research, Innovation and the Challenges of High-Quality Education*; Rienties, B., Hampel, R., Scanlon, E., Whitelock, D., Eds.; Routledge: London, UK, 2022.
16. Brajnik, G.; Yesilada, Y.; Harper, S. Is accessibility conformance an elusive property? A study of validity and reliability of WCAG 2.0. *ACM Trans. Access. Comput. TACCESS* **2012**, *4*, 1–28. [CrossRef]
17. Zhang, X.; Tlili, A.; Nascimbeni, F.; Burgos, D.; Huang, R.; Chang, T.W.; Jemni, M.; Khribi, M.K. Accessibility within open educational resources and practices for disabled learners: A systematic literature review. *Smart Learn. Environ.* **2020**, *7*, 1–19. [CrossRef]
18. Hall, T.E.; Cohen, N.; Vue, G.; Ganley, P. Addressing learning disabilities with UDL and technology: Strategic reader. *Learn. Disabil. Q.* **2015**, *38*, 72–83. [CrossRef]
19. Rao, K.; Ok, M.W.; Smith, S.J.; Evmenova, A.S.; Edyburn, D. Validation of the UDL reporting criteria with extant UDL research. *Remedial Spec. Educ.* **2020**, *41*, 219–230. [CrossRef]
20. CAST. Top 5 UDL Tips for Fostering Expert Learners. 2017. Available online: http://castprofessionallearning.org/wp-content/uploads/2017/08/cast-5-expert-learners-1.pdf (accessed on 1 September 2022).
21. Handoko, E.; Gronseth, S.L.; McNeil, S.G.; Bonk, C.J.; Robin, B.R. Goal setting and MOOC completion: A study on the role of self-regulated learning in student performance in massive open online courses. *Int. Rev. Res. Open Distrib. Learn.* **2019**, *20*. [CrossRef]
22. Iniesto, F.; Hillaire, G. UDL and its implications in MOOC accessibility evaluation. In *Open World Learning: Research, Innovation and the Challenges of High-Quality Education*; Rienties, B., Hampel, R., Scanlon, E., Whitelock, D., Eds.; Routledge: London, UK, 2022.
23. Rodríguez-Ascaso, A.; Letón-Molina, E. Materiales Digitales Accesibles. 2018. Available online: http://e-spacio.uned.es/fez/view/bibliuned:EditorialUNED-aa-EDU-Arodriguez-0003 (accessed on 1 September 2022).
24. Rodrigo, C.; Iniesto, F.; García-Serrano, A. Reflections on Instructional Design Guidelines From the MOOCification of Distance Education: A Case Study of a Course on Design for All. In *UXD and UCD Approaches for Accessible Education*; IGI Global: Hershey, USA, 2020; pp. 21–37.
25. Myers, K.K.; Powers, S.R. Mixed methods. In *The International Encyclopedia of Organizational Communication*; Wiley: New York, NY, USA, 2017; pp. 1–11.
26. Landis, J.R.; Koch, G.G. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* **1977**, 363–374. [CrossRef]
27. Gavin, H. Thematic analysis. In *Understanding Research Methods and Statistics in Psychology*; SAGE: London, UK, 2008; pp. 273–282.
28. Durham, M.F.; Knight, J.K.; Couch, B.A. Measurement Instrument for Scientific Teaching (MIST): A tool to measure the frequencies of research-based teaching practices in undergraduate science courses. *CBE—Life Sci. Educ.* **2017**, *16*, ar67. [CrossRef] [PubMed]
29. Iniesto, F.; Rodrigo, C.; Hillaire, G. *Applying UDL Principles in an Inclusive Design Project Based on MOOCs Reviews*; Routledge: New York, NY, USA, 2019.
30. Chen, C.H.; Tsai, C.C. In-service teachers' conceptions of mobile technology-integrated instruction: Tendency towards student-centered learning. *Comput. Educ.* **2021**, *170*, 104224. [CrossRef]
31. Poquet, O.; Dawson, S.; Dowell, N. How effective is your facilitation? Group-level analytics of MOOC forums. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Vancouver, BC, Canada, 13–17 March 2017; pp. 208–217.