Technology, Knowledge and Learning (2026) 31:1039–1069
https://doi.org/10.1007/s10758-025-09858-2

ORIGINAL RESEARCH

Improving Learning Through Evaluative Judgement and Feedback Using a Technology-Enhanced Assessment Environment

María Soledad Ibarra-Sáiz[1,2] · Miguel Ángel Gómez-Ruiz[1,2] · Antonio Balderas[3] · Gregorio Rodríguez-Gómez[1,2]

Accepted: 28 April 2025 / Published online: 26 May 2025
© The Author(s) 2025

Abstract

It is essential for students to develop their evaluative judgement in post-secondary education. This refers to their ability to weigh up and make decisions regarding the quality of their work and that of others. It is also necessary to facilitate the feedback process as an aspect which supports learning. This paper presents the evaluation of EvalCOMIX-FLOASS, a technological environment developed to facilitate monitoring and assessment of learning outcomes, enhancing evaluative judgement and feedback. This technological environment has been supported within the assessment-as-learning approach and developed to facilitate monitoring and assessment of learning outcomes. The evaluation process used a mixed methodology with a multi-phase evaluative design to gather the perceptions of 275 students and 43 teachers who interacted with this technological environment on 39,093 occasions. It was rated very positively for facilitating development of evaluative judgement and the feedback processes. In the students' opinion, this is an environment which helps them rate their learning and the strategies that they use to guide this process, plus the mechanisms required to make progress. In the teachers' opinion, this is a useful environment which allows them to monitor and assess achievement levels from the learning outcomes. Future lines of research are suggested, requiring a design rethink for this type of technological environment so that it might facilitate innovative assessment processes, taking into consideration the opportunities and threats represented by artificial intelligence.

Keywords Assessment as learning · Assessment methods · Evaluative judgement · Feedback · Technology-enhanced assessment · Technological environment

Abbreviations

CATEF_FLOASS_E	Environment technological acceptance questionnaire for students
CATEF_FLOASS_P	Environment technological acceptance questionnaire for teachers
DF	Design-facility of use
EJ	Evaluative judgement
GE	Generalisation
FE	Feedback

Extended author information available on the last page of the article

LA	Learning analytics
LMS	Learning management system
LO	Learning outcomes
M	Mean
Mdn	Median
MISCA	Message implementation student context agents
PERMANOVA	Permutational multivariate analysis of variance
PLS-SEM	Partial least squares structural equation modeling
SA	Satisfaction
SD	Standard deviation
SPARK-PLUS	Software for flexible self & peer assessment and feedback (Sydney University)

1 Introduction

The purpose of assessment as learning (Yan & Boud, 2022) is to act as a guide and provide opportunities for the student to be able to monitor and think critically about their own learning, thereby identifying the steps they should follow. Consequently, technology can play an important role in enhancing learning insofar as it eases development of evaluative judgement and improves feedback. In the post-secondary education context, there is a clear need for students to develop their own evaluative judgement, meaning the ability to rate and make decisions regarding the quality of their own and other people's work (Carless & Boud, 2018; Tai et al., 2018). This ability, along with self-knowledge and the perception of their own efficacy, are fundamental elements that contribute to human intelligence and make it stand out from work performed by machines, in a world evermore influenced by the development of artificial intelligence (Bearman & Luckin, 2020; Bearman et al., 2024).

In turn, feedback is one of the critical, central and necessary components that relate assessment to learning. Currently, feedback is conceived "as a process through which learners make sense of information from various sources and use it to enhance their work or learning strategies" (Carless & Boud, 2018, p. 1315). This conceptualisation implies accepting that feedback is not simply information passed on by the teacher, but a process or a series of processes where the student is an active participant who must construct the meaning of the contributions received so that this information might have a real impact on their present and future learning (Henderson et al., 2019).

This study presents the EvalCOMIX-FLOASS technology environment and the results of its evaluation. This technology is based on the assessment-as-learning approach, developed to monitor and assess learning outcomes as a tool which facilitates evaluative judgement and feedback.

2 Theoretical Background and Research Questions

2.1 The Importance of Developing Evaluative Judgement and Its Practice Among Post-secondary Education Students

Use of educational technologies, which coherently enhance evaluative judgement among students, must match practical strategies that work on and consolidate this ability, grounded

in the epistemological progress made in this field. These strategies, described below, are illustrated in Fig. 1.

To make sure that the students' assessments and contributions are coherent, unbiased and fair, it is fundamental that they have significant skills in evaluative practices (Ibarra-Sáiz et al., 2020). In the same way, this student decision-making, which involves metacognitive processes, must be refined using contributions from other people, not only regarding the quality of the actual work, but also the ability to judge for themselves. The role of feedback is thereby essential in this aspect (Henderson et al., 2019).

Sufficient time, practice and implication from the students when assessing any phase of the process (planning, development and qualification) plus well-determined criteria and transparency in the assessment processes, making them public, accessible and comprehensible, are all practices that might particularly boost students' evaluative judgement (Ibarra-Sáiz & Rodríguez-Gómez, 2019).

Improving students' evaluative judgement is also supported by using explicit instruments such as rubrics, which increase the transparency of assessment, reduce anxiety, support feedback, and increase self-efficacy and self-regulation (Panadero & Johnson, 2013). In addition, the use of exemplifications with quality standards in tasks is crucial. Technology, along with participatory methods (self-assessment, peer assessment, co-assessment), bridges theory and practice in developing this skill (Khosravi et al., 2021).

Continuous student participation in evaluative practices in the university environment, plus dialogical interactions regarding their performance, encourage feedback and self-feedback (Panadero et al., 2019a) and are tools which can encourage self-regulated learning, improve future work and even lead to higher scores for the students, focussing on achieving sustainable feedback (Carless et al., 2011).

In the same way, there is clearly a need to help the student become assessment-literate to guarantee their reasoned, effective and complete implication in the assessment processes (Smith et al., 2013). Consequently, and due to its importance in developing self-regulation strategies that can be used to monitor and guide learning actions, both in current and novel



Fig. 1 Ten ways of boosting students' evaluative judgement

situations, it might be considered that learning how to evaluate should become a fundamental part of all study plans (Panadero et al., 2019b).

To conclude, fundamental components to develop evaluative judgement should comprise (Boud et al., 2018a, 2018b): in-depth comprehension of the training context where the task is being developed, assimilation of the quality standards required in the work, specific assessment criteria used to evaluate the task and the need for evaluative experience using participatory methods and thinking and guidance in this regard.

2.2 Feedback as a Basis for Learning from and for Assessment

The information obtained by the students from various sources must come from high quality timely contributions on their performance in the appropriate form to be able to extract all educational potential (Ajjawi & Boud, 2018; Hattie & Timperley, 2007). In the same way, the feedback concept should be developed to include the participants, their contextual variables and the supports in which they are provided (Gómez-Ruiz & Quesada-Serra, 2020; Ruelmann et al., 2023).

No longer is only the teacher named as the expert and person responsible for providing the students with the information, rather more a multi-directional process must be considered where students provide information to other classmates, to themselves and also to the teacher (López-Pastor & Sicilia-Camacho, 2017; Nicol et al., 2014). The fact that the students take centre stage in the feedback actually encourages learning among both the recipients and issuers of the information, as it enhances their ability to argue and put across ideas, as well as the actual evaluative judgement mentioned above (Ibarra-Sáiz et al., 2020).

Likewise, the implication of the students' sociocultural context should also be explored, paying attention to the emotional impact of the feedback processes, and considering intrapersonal and interpersonal factors to determine the real and productive impact it might have on students' learning (Henderson et al., 2019).

It is important to outline the educational focus of the feedback by using related tasks (Morrell, 2021) to boost chances that the students might use the information and maximise their learning (Boud & Molloy, 2013), and also support self-regulation among these students (Afzaal et al., 2024; Nicol & Macfarlane-Dick, 2006; Wolterinck-Broekhuis et al., 2024). It is therefore essential that the students take part in subsequent related endeavours to be able to try out their new understanding and begin the cycle again (Henderson et al., 2019). This also encourages feedback that might benefit generalisation to other contexts, not only academic but also professional, known as *feedforward* (Sadler, 2010; Tan & Chen, 2022).

In relation to and regarding analysis of 14 feedback models, Panadero and Lipnevich (2022) propose the MISCA model (Message Implementation Student Context Agents). This revolves around considering the students' individual characteristics to adapt the process to their specific needs. Regarding the message, they emphasise that it must be high quality, meaning that it must be clear, accessible and usable. During implementation, it should be understood that feedback influences self-regulation and the students' emotional and motivational processes. Regarding the context, the focus should be on the general educational design, the atmosphere in the classroom, and also the institutional context and the actual educational system. Finally, in terms of agents, not only the teacher but also classmates and the actual learner can create feedback.

In conclusion, we must set ourselves new challenges and perspectives for the future regarding personalisation of feedback (Deeva et al., 2021) working from the, as yet uncertain, development of artificial intelligence, which might bring about changes in education, assessment and specifically in the feedback. It might be possible to improve efficiency and personalisation of feedback, but this should be treated critically (Zawacki-Richter et al., 2019).

2.3 Technologies that Enhance Assessment and Learning

Over the last few years, use of learning analytics (LA) has become more relevant in the educational context (Peña-Ayala, 2017; Berk et al., 2022), although the European-level study carried out by Tsai et al. (2020b) confirms, firstly, that there is little empirical evidence of students' implication in the use of LA to develop their self-regulated learning skills and, secondly, the need to justify and ground the LA in the learning sciences and actively include students as an interest group when designing and implementing the LA.

Using technologies to facilitate feedback processes has encouraged multiple realities, not only as a dialogical practice or in writing, but that the inclusion of audio, video, screencast or annotations on the actual work, among others, are effective methods to add value to in-person and traditional environments, encouraging personalisation and comprehension of the information in mass educational fields (Rodríguez et al., 2022; Pitt & Winstone, 2020; Rönn & Pettersson, 2023). Consequently, use of technology might be considered an opportunity to make feedback more effective and even be able to automate creation and transfer of information to increase individual contributions received by each student (Jensen et al., 2021), make technology more effective in terms of improving writing and joint work (Zou et al., 2023) or the use of LA to provide personalised feedback (Kärner et al., 2021; Pardo et al., 2019).

Despite increased use of technology in assessment (Farrell & Rushby, 2016), little research has been produced on the role of these technologies in configuring wider pedagogic approaches in post-secondary education, as any new and innovative approach in assessment implies a new challenge to understand how technologies are being used (Caird & Lane, 2015). These authors highlight the importance and need to conceptualise the complex and changing role of technologies in post-secondary education teaching models. When lacking, it is difficult to weigh up the benefits and limitations of using technologies to support the learning. In this respect, Bearman et al. (2020) outline the need to focus activity on three necessary reinventions: (1) rework the assessment *status quo* to improve the assessment process design, making use of these technologies; (2) explore new assessment designs that consider the digital world; and (3) reinvent assessment designs for new and emerging forms of knowledge.

Educational institutions provide teachers with on-line tools which can facilitate their students' learning and help them throughout the assessment process (Balderas et al., 2018). Specifically, these tools are included in the virtual learning management systems (LMS) to subsequently facilitate the use of LA (Verdú et al., 2021), and comprise myriad tools which are likely to be used. In this respect, Moodle has a Workshop tool that allows students to hand in any digital content which might be subsequently reviewed by classmates or the students themselves according to an assessment form defined by the teacher (Deepak, 2017). However, although a powerful tool with multiple settings that teaching staff can adapt to their classes' needs, the Moodle Workshop has some limitations, mainly referring to exporting data and visualising results (Chaparro-Peláez et al., 2019).

These technological tools have been evaluated with reference to the principles of user experience (UX), which has a long tradition and various theoretical models associated with it (International Organization for Standardization, 2019). Drawing on this tradition, this article analyses widely recognised variables, such as facility of use (Davis, 1989; Venkatesh et al., 2003), generalisation and satisfaction (Hassenzahl, 2008), and introduces two new variables which are extremely important in the field of educational assessment and must be taken into consideration when analysing any technological tool that seeks to improve evaluation and learning in an educational context: feedback (Nicol & Mcfarland-Dick, 2006) and evaluative judgement (Boud et al., 2018a, 2018b).

2.4 The EvalCOMIX-FLOASS Environment

The EvalCOMIX-FLOASS environment is a web service integrated into Moodle as a block. Working from a previous design, within Moodle, of evaluative activities comprising the assessment and learning tasks, in line with the assessment-as-learning approach, the EvalCOMIX-FLOASS environment is intuitive and easy to use to (Fig. 2):

1. Manage each subject's learning outcomes (LO) and skills.
2. Devise and manage various assessment instruments (checklist, assessment scale, rubric, etc.).
3. Connect learning outcomes and skills that are assessed in each dimension of the instruments.
4. Plan assessments by linking each instrument to the activities, programming the assessment methods, choosing between individual or group assessments and choosing to include all the scores or only the mid-range scores (within a specific threshold).
5. Obtain reports for each student, work groups and for all the students regarding how far learning outcomes and skill development have been achieved.

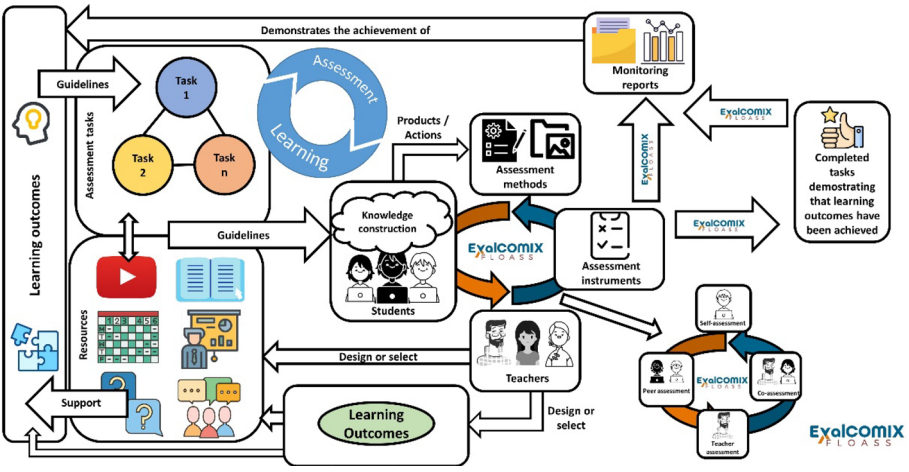


Fig. 2 Components of the EvalCOMIX-FLOASS technological environment (Rodríguez-Gómez et al., 2022)

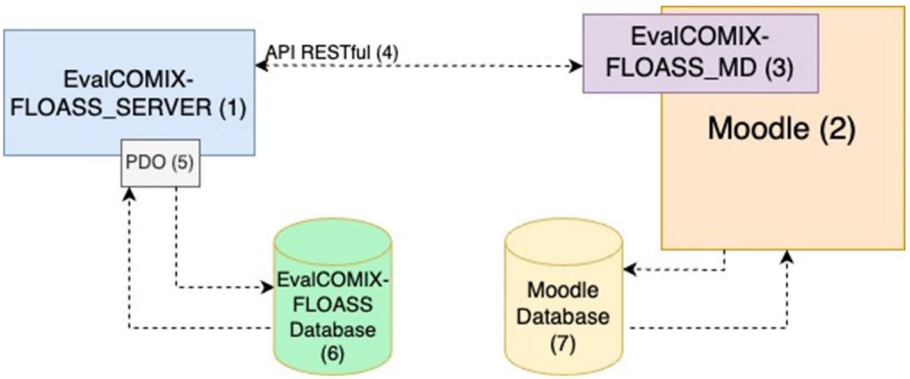


Fig. 3 Architecture of the EvalCOMIX-FLOASS technological environment

The architecture of the EvalCOMIX-FLOASS environment (Fig. 3) is based on three fundamental components: EvalCOMIX-FLOASS_Server (1), Moodle (2) and EvalCOMIX-FLOASS_MD (3).

EvalCOMIX-FLOASS_Server (1) is a web service initially developed as a tool to create and deploy assessment instruments in virtual-course environments (Ibarra-Sáiz & Rodríguez-Gómez, 2020), but which has expanded its functional features to include management of learning outcomes and skills associated with these courses. This is a web application that implements a RESTful API (4), developed in PHP programming language with the Symfony framework. To store information, EvalCOMIX-FLOASS_Server uses a PDO interface (5) with specific controllers for MySQL and PostgreSQL, and thereby connects with its database (6). EvalCOMIX-FLOASS_Server is continuously being developed, and the current version is 4.3.3.[1]

Moodle (2) implements its own database (7), depicted separately in the diagram of the EvalCOMIX-FLOASS environment (6).

EvalCOMIX-FLOASS_MD (3) is a plugin to integrate EvalCOMIX-FLOASS_Server on the Moodle learning platform. EvalCOMIX-FLOASS_MD was developed with the PHP programming language. As it is a plugin for integration with EvalCOMIX-FLOASS_Server, this uses its RESTful API (4). The current version of EvalCOMIX-FLOASS_MD is 4.4.1.[2]

It is important to highlight that the EvalCOMIX-FLOASS environment has been successfully implemented in more than 40 courses at seven Spanish universities, demonstrating its reliability and maturity in real-life educational environments. The system has achieved a Technology Readiness Level (TRL) of 9 (Salazar & Russi-Vigoya, 2021), indicating its readiness for widespread deployment. This high level of maturity is further supported by the ISO/IEC 25000 functional suitability and maintainability certifications issued by AENOR, which ensure that the system complies with international standards for software quality.[3]

[1] https://evalcomix-floass.uca.es/en/home/.

[2] https://moodle.org/plugins/block_evalcomix.

[3] https://iso25000.com/index.php/en/news/274-universidad-de-cadiz-obtains-the-iso-iec-25000-maintainability-certificate-for-their-software-evalcomix-floass-server-and-evalcomix-floass-md.

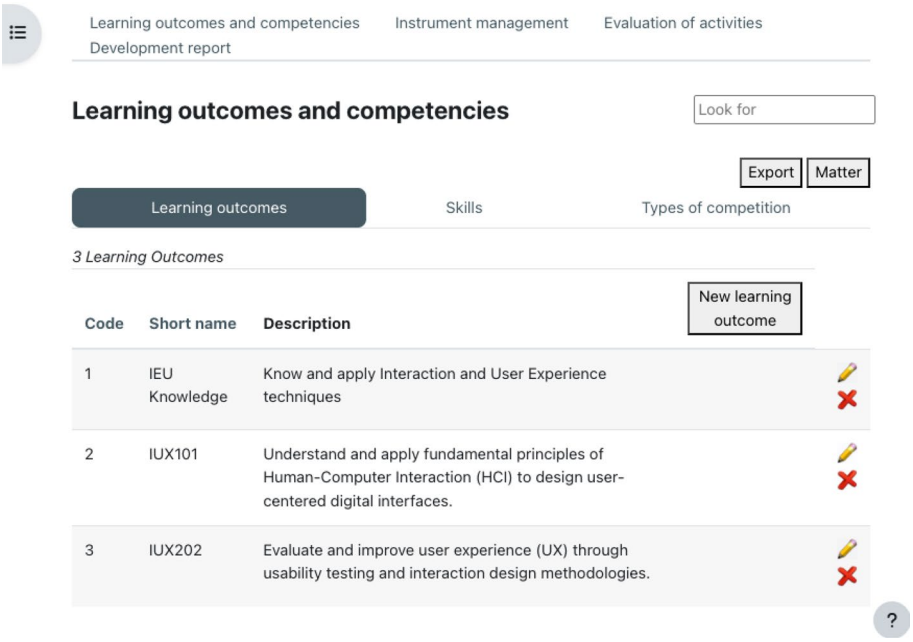


Fig. 4 Defining learning outcomes with EvalCOMIX-FLOASS

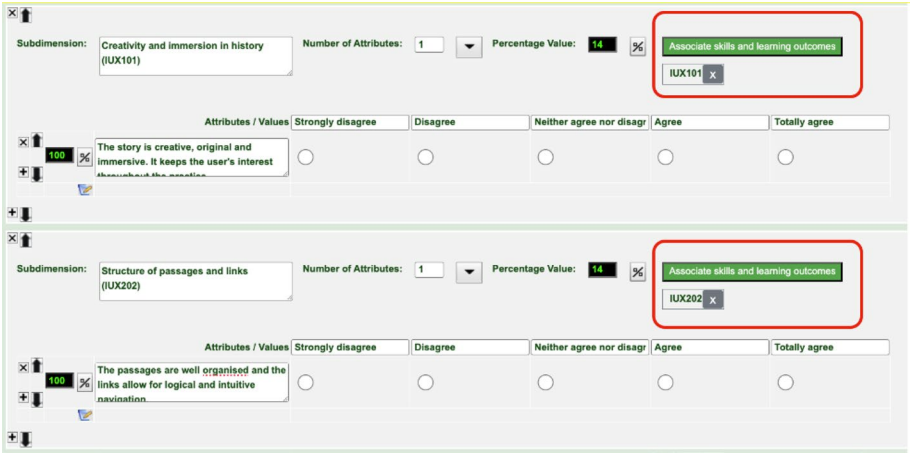


Fig. 5 Associating indicators from assessment instruments with learning outcomes

Various other LMS also provide self-assessment and peer assessment tools. Boasting a more modern design than Moodle, Canvas (Sullivan & Watson, 2015) includes a powerful tool which is comparable to Workshop by Moodle, although its interface is more up-to-date and user-friendly for students (Kobayashi, 2020). However, the distinctive advantage of EvalCOMIX-FLOASS is its capacity to map out all these activities with the skills and learning outcomes from the course, making it more significant. For instance,

EvalCOMIX-FLOASS allows direct alignment of activities with learning outcomes and competencies, ensuring meaningful learning experiences. Figure 4 shows the screen where learning outcomes are defined in the EvalCOMIX-FLOASS plugin, while Fig. 5 shows the section where items from an assessment instrument are linked to competences and learning outcomes. Its seamless integration as a Moodle plugin eliminates the need for external tools, enhancing usability and ease of adoption. Another LMS such as Sakai (Petersen & Groenewald, 2021) also offers this type of assessment, although it presents usability problems for the students, an issue that EvalCOMIX-FLOASS has addressed and that will be assessed in later sections.

Therefore, EvalCOMIX-FLOASS offers a series of advantages when compared with other available tools. Its integration with Moodle allows the use of various personalised evaluation instruments in every activity available in this LMS. It also provides direct access and data synchronisation, simplifying the workflow of teachers and students—a feature that functions less effectively in other tools. With regard to the quality of its feedback mechanisms, EvalCOMIX-FLOASS allows richer, more personalised feedback to be delivered, with the option to add detailed comments in the evaluation instruments and obtain partial and overall results and all the information collected during the evaluation process instantaneously.

2.5 Objective and Research Questions

The study objective focuses on validating the EvalCOMIX-FLOASS environment, by answering the following research questions:

- CI1 To what degree do teachers and students technologically accept the EvalCOMIX-FLOASS environment? Do the perceptions of these two collectives differ?
- CI2 Is user satisfaction with the EvalCOMIX-FLOASS environment related to evaluative judgement and feedback?

This study presents several significant innovations in the field of educational assessment. Firstly, EvalCOMIX-FLOASS is a recently-developed tool that assists in-depth integration of assessments with the skills and learning outcomes (Sect. 3.3). The mixed methodology used in this study, combining quantitative and qualitative focus points (Sect. 3.2), provides an exhaustive evaluation of the tool in diverse educational contexts. Furthermore, the agile, group-consensus approach to develop EvalCOMIX-FLOASS is a novel way of designing educational technological tools (Sect. 3.2). This study not only validates the tool empirically using questionnaires and focus groups (Sect. 3.4), but it also offers a framework for future research on its implementation and effectiveness.

3 Context and Methodology

3.1 Design

This used a mixed methodology, specified in a multistage evaluative design (Creswell, 2015; Decuir-Gunby & Schutz, 2017), which combines quantitative and qualitative approaches (Fig. 6). The first phase, which is qualitative, focused on developing the EvalCOMIX-FLOASS environment, and began with the literature review and analysis of

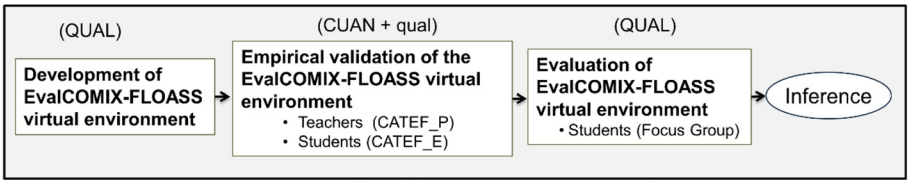


Fig. 6 Multistage evaluation design

existing tools such as Doubtfire (Cain et al., 2020) and SPARK-PLUS (Willey et al., 2009), using an agile, group-consensus methodology (Johnson & Morgan, 2016). The empirical evaluation phase was performed using a survey process among teaching staff and students who use the environment in their daily academic activity. In the third phase, use of the EvalCOMIX-FLOASS environment was evaluated using the students' rating, expressed via a focus group.

The mixed methods approach was chosen to provide a comprehensive understanding of the EvalCOMIX-FLOASS environment. This article presents the main outcomes of phases two and three, where qualitative data were used to substantiate the quantitative findings.

3.2 Evaluation Questionnaires

Two instruments were designed to empirically validate the EvalCOMIX-FLOASS environment: *CATEF_FLOASS_P environment technological acceptance questionnaire* intended for teachers and *CATEF_FLOASS_E environment technological acceptance questionnaire* for students. These questionnaires are based on linear combinations of the study variables (Henseler, 2021). Cognitive and behavioural aspects played an essential role in the evaluation of each questionnaire item by the respondents, so these items were used to build a formative index (Hair et al., 2022). Comprehensive indicators for formatively measured constructs should be identified using a rigorous qualitative approach. Consequently, the content validation for these questionnaires followed an iterative process among five judges using the group consensus method (Johnson & Morgan, 2016) to review the various constructs and indicators, specified in three iterations, until an agreement was reached on the latest version of the instruments. The two instruments are organised into five common dimensions: Design-Facility of use (DF), Feedback (FE), Evaluative judgement (EJ), Satisfaction (SA) and Generalisation (GE) that include 18 Likert-type questions (0–5) for the teachers and 14 for the students, in this case completed by two open questions. It took five minutes for each collective to complete the questionnaire.

The internal consistency perspective underpinning reflective measurement model evaluation cannot be applied to formative models since formative measures do not necessarily covary (Hair et al., 2022). Furthermore, formative indicators are assumed to be error free (Diamantopoulos, 2006), which means that the internal consistency reliability concept is not appropriate. Consequently, the Partial Least Squares Structural Equation Modelling (PLS-SEM) method was performed. With values of variance inflation factor (VIF) between 1.08 and 2.10 for teachers and 1.00 and 2.11 for students, we can conclude that collinearity does not reach critical levels in any of the formative constructs and is not an issue for the estimation of the PLS path model (threshold value of 5). Adjustment measures of SRMR = .09 (teachers) and SRMR = .05 (students) were obtained.

The questionnaires were applied using the SurveyMonkey platform for a limited period of time after completion of the subjects in which the EvalCOMIX-FLOASS environment was implemented in the first semester of the 2022–23 academic year.

The CATEF_E questionnaire was applied in 15 subjects of which seven (46.67%) belong to three undergraduate degree courses (two from Education and one from Economics), and eight (53.33%) to two master's degree courses (one in Education and another in Engineering). Five of the master's degree subjects correspond to an online master's degree and the other ten to classroom-based courses. In the third phase, student evaluation was considered in greater depth using a focus group of students from a master's degree that used the EvalCOMIX-FLOASS environment in all subjects. These students took subjects developed from the assessment-as-learning perspective and their teachers had at least four years' prior experience using the EvalCOMIX plugin (previous version of the EvalCOMIX-FLOASS environment). The focus group was made up of five students, 80% women and 20% men.

3.3 Sample

In the empirical evaluation phase, the sample comprised 43 teachers and 275 students who participated in the study. The sampling was intentional. To take part, the teachers had to meet one of the following requirements: have taken the "Design and assessment of learning outcomes" training workshop or have used the EvalCOMIX-FLOASS environment in at least one subject. As for the students, they were required to study one of the subjects

Table 1 Demographic characteristics of the teacher sample

Characteristics		f	%
Age	Under 30 years old	3	7
	Between 30 and 39 years old	7	16.3
	Between 40 and 49 years old	7	16.3
	Between 50 and 59 years old	22	51.2
	60 or over	4	9.3
Gender	Female	25	58.1
	Male	16	37.2
	Others	2	4.7
University[a]	UCA	15	34.9
	UDC	6	14
	UCM	4	9.3
	UNIOVI	4	9.3
	UPV/EHU	4	9.3
	URV	5	11.6
	Others	5	11.6
Knowledge branches	Social Sciences	40	93
	Engineering and Architecture	2	4.7
	Art and Humanities	1	2.3
Total		43	100

[a]UCA: Universidad de Cádiz, UDC: Universidad de A Coruña, UCM: Universidad Complutense de Madrid, UNIOVI: Universidad de Oviedo, UPV/EHU: Universidad del País Vasco, URV: Universidad Rovira i Virgili

Table 2 Demographic characteristics of the student sample

Characteristics		f	%
Age	21 years old or under	154	56
	Between 22 and 25 years old	94	34.2
	26 or over	27	9.8
Gender	Female	195	70.9
	Male	80	29.1
Knowledge area	Economics	74	26.9
	Education	197	71.6
	Engineering	4	1.5
Type of teaching	Degree	225	81.8
	Master's degree	50	18.2
Total		275	100

given by the teachers from the sample. The characteristics of both samples are presented in Tables 1 and 2.

The sample from this study mainly comprises Social Science teachers because the teachers from this area have demonstrated greater interest in joining the project from the outset. However, more teachers from other disciplines are gradually being included, thereby broadening the diversity of the represented areas.

Several teachers in the study have prior experience working with Workshop by Moodle, which is known to be easy to use, and included as a further activity within Moodle. Furthermore, the teachers taking part in this study have experience in a wide range of projects on evaluating skills and learning outcomes, which backs up their observations and recommendations. In turn, the students use the Moodle environment in their various degree programmes.

All the students came from the University of Cadiz. Their mean age was 21.98 (SD = 3.98) years old, and the mean age in the degree courses was four years younger (M = 21.2, SD = 3.47) than in the master's courses (M = 25.2, SD = 4.4). The CATEF_E questionnaire was applied in 15 subjects of which seven (46.67%) belong to three undergraduate degree courses, two from Education and one from Economics, and eight (53.33%) to two master's degree courses in Education and one in Engineering. Five of the master's degree subjects correspond to an online master's degree and the other ten subjects to classroom-based courses.

Although the participating teachers originate from more than six different universities, the students are all from the same university. This decision was taken mainly for logistical and coordination reasons, as the main researchers for the project belong to this institution. Furthermore, our university ethics committee only authorises working with students from our own institution. This configuration makes it easier to manage the study without compromising the validity of the results, as the main goal is to assess how effective the tool might be through the feedback provided by the teachers from various academic institutions.

In the third phase, in an attempt to dig deeper into the assessment, information was compiled, using a focus group, on the master's degree students that used the EvalCOMIX-FLOASS environment in all their subjects. These subjects were developed from the assessment-as-learning perspective and their teachers had prior experience using

Table 3 Summary of interactions and assessments done by the teachers and students with the EvalCOMIX-FLOASS environment

	Interactions			Assessments		
	Teachers (average)	Students (average)	Total	Teachers (average)	Students (average)	Total
Degree	3,018 (215.6)	20,306 (58.5)	23,324	1,496 (106.9)	10,464 (30.2)	11,960
Master	3,258 (130.3)	12,511 (69.9)	15,769	1,103 (44.1)	7,933 (44.3)	9,036
Total	6,276	32,817	39,093	2,599	18,397	20,996

the EvalCOMIX plugin (previous version of the EvalCOMIX-FLOASS environment) which had been implemented in their subjects for at least four years. There were 5 students in the focus group, 80% women and 20% men.

3.4 Data Analysis

3.4.1 Quantitative Data Analysis

The quantitative analysis involved several statistical procedures. Firstly, trend and dispersion analysis was run using descriptive statistics (Md, M and SD) and box graphics in both questionnaires' dimensions and items. Subsequently, answers to the student questionnaires were analysed by level, gender and age; and teacher questionnaires depending on the university, gender and age, using the PERMANOVA (Anderson, 2017) and Kruskall-Wallis tests. Comparisons were also established in the answers from both collectives. Finally, the Partial Least Squares Structural Equation Modelling technique was used (PLS- SEM). For this analysis, R (R Core Team, 2021) and SmartPLS 4 (Ringle et al., 2022) were used.

3.4.2 Qualitative Data Analysis

The qualitative part of the study used an open and axial coding focus, following the qualitative data analysis methodology (Flick, 2014; Miles & Huberman, 1994; Rodríguez-Gómez et al, 1999). The sampling strategy was intentional, selecting participants from the focus group based on their experience with the tool and their willingness to provide detailed feedback. The data was collected through semi-structured interviews and discussions among the focus group, making it possible to explore students' perceptions and experiences in depth. The qualitative analysis involved several stages: First, the transcriptions of the interviews and focus group discussions were coded using a system of mixed categories. Thematic analysis was then performed to identify recurring patterns and topics. The quantitative data are complemented by the qualitative results, providing a comprehensive overview of the extent to which EvalCOMIX-FLOASS is effective and accepted.

3.4.3 Validation of Results

Finally, to ensure the validity and reliability of the qualitative results, the data was triangulated by comparing the qualitative and quantitative findings and seeking coherence and discrepancies. This triangulation strengthened the robustness of the study's conclusions.

4 Results

4.1 Intensity of Use of the EvalCOMIX-FLOASS Environment

Table 3 presents the interactions and assessments by the teachers and students who interacted with the environment when teaching subjects within the framework of the empirical tests for the EvalCOMIX-FLOASS environment. A total of 39,093 interactions were registered with the environment (consulting, visualising, etc.), of which 53.7% (20,996) were assessments performed using an assessment instrument for self-assessment, teacher assessment or peer assessment.

4.2 Technological Acceptance and Differences Between Students and Teachers

4.2.1 The Students' Perception

The results from Table 4 demonstrate a high degree of student satisfaction regarding the EvalCOMIX-FLOASS environment. Aspects were evaluated relating to design and facility of use (DF), feedback (FE), evaluative judgement (EJ) and general satisfaction (SA), obtaining average values in all the aspects and dimensions that were close to or over 4, indicating a very positive overall perception. On the other hand, Table 5 shows the main (frequency ≥ 5) positive aspects and proposed improvements provided by the students taking part in the questionnaire and in the focus group. The quantitative and qualitative results corresponding to both evaluations are mentioned and compared below.

Firstly, we should highlight the positive student outlook concerning the environment's ease of use (DF), where 83.3% and 85.4% of students agree or completely agree that the environment is simple and intuitive (DF2) and that it is easy to use with minimum instruction (DF3), respectively. These quantitative results are backed up by the students' qualitative comments. In this case, most students (71 answers) praise the tool as simple, fast and intuitive to use. For example, one student mentioned in the questionnaire:

> "I might highlight its great simplicity, and how intuitive the tool is. It might be an option to consider if you are looking for speed and ease" (C159).

However, the improvements proposed by the students in this field have varied widely, covering a total of 18 different subcategories. The most significant included the need to improve the environment's design and aesthetic (35), the assessment saving function (12) and how it was sometimes sluggish when interacting with the assessment instruments (8). As examples, the students consider the following:

> "The most emphasis should be on is its actual design. It seems a little 'old-fashioned' technologically speaking" (C129).

This perception of possibilities for improvement matches the quantitative results compiled for the DF1 dimension, where 62.9% of students consider that the environment has an attractive design. The remaining dimensions being evaluated are deemed to be over 74%.

As far as feedback (FE) is concerned, the students highlight how important it is for them that the assessment results are available quickly (FE2) plus the chance to receive information on developing skills and learning outcomes (FE3), with 88.7% and 83.7% respectively who agree and completely agree. Qualitatively, it should be mentioned that the students

Table 4 Central trend and dispersion measurements in the dimensions and items in the CATEF_E questionnaire

Dimensions and items	Mdn	M	SD
Design-Facility of use (DF)	*4.33*	*4.13*	*.82*
DF1 It has an attractive design	4.00	3.70	1.19
DF2 It is simple and intuitive	5.00	4.27	.94
DF3 It can be used with minimal instructions	5.00	4.41	.86
Feedback (FE)	*4.25*	*4.18*	*.69*
FE1 I receive information in each of the dimensions and its aspects that I can use to improve my future tasks	4.00	4.16	.86
FE2 I get the assessment results quickly	5.00	4.41	.87
FE3 I can find out information on developing my skills and learning outcomes	4.00	4.16	.87
FE4 The reports generated by EvalCOMIX-FLOASS are interesting and I am satisfied with them	4.00	4.00	1.00
Evaluative judgement (EJ)	*4.32*	*4.13*	*.68*
EJ1 I can find out about the assessment instruments beforehand plus the criteria and aspects that are going to be rated	4.00	4.16	.94
EJ2 I understand the assessment process	5.00	4.32	.89
EJ3. It drives and encourages my participation in the assessment	5.00	4.42	.78
EJ4 It develops my evaluating skill	5.00	4.40	.90
Satisfaction (SA)	*4.25*	*4.14*	*.84*
SA1 I am happy to use the EvalCOMIX-FLOASS tools to assess my tasks	4.00	4.21	1.03
SA2 On a scale of 0 to 10, indicate your overall satisfaction with the EvalCOMIX-FLOASS tool	8.00	8.16	1.54
Generalisation (GE)			
SA1 I would like to see more general use of EvalCOMIX-FLOASS across other subjects at my university	5.00	4.12	1.20

Table 5 Main positive aspects and proposed improvements provided by the students

Category	Positive aspects	Proposed improvements
Design and Facility of use (DF)	It is a simple, fast and intuitive tool (71)	General design and aesthetic of the interface (35)
		Assessment saving function (12)
		Interface is slow to use (8)
		Buttons to perform the assessment should be more separate and clearer (5)
		Ordering of the list and search for students (5)
Feedback (FE)	Get the assessment results and comments quickly(38)	Slow to generate reports and their processing (8)
	Receiving information from each dimension and the aspects to be improved (19)	Information should be available for the whole work group (5)
	Allows peer assessment and feedback (11)	
	There is the chance to add comments (6)	
Evaluative judgement (EJ)	It encourages student participation in the assessment (29)	
	Finding out about the assessment instruments with the criteria and aspects that are going to be assessed (21)	
	It develops students' assessment skills (14)	
No aspect		Nothing requires improvement (21)

particularly appreciated that this tool gave them quick access both to the results of their assessment and the comments provided (38). In the same way, 19 mentioned that this environment provided them with accurate information on each dimension being evaluated, and clearly stated aspects that they must improve in future projects. Regarding these questions, they mentioned the following:

> "Getting your scores immediately, being able to read each evaluator's opinion on your work" (C24).
> "I think it's also quite useful when they give you a grade or a score to know why you got that score, because, for example, I get a seven, but why did I get a seven? Where did I go wrong? If I go to the assessment rubric or scale, I can see where I made the most mistakes or where I did best, and that gives students information that they can take into consideration on future occasions." (GF5)

The most outstanding positive aspects are completed by the relevance of being able to carry out peer assessments, encouraging feedback among classmates (11 answers), and that it is possible to add comments in the various instruments (6).

> "It's simple and easy to use, I think it's an ingenious idea to have students participate in evaluations, self-evaluations and peer evaluations and be able to access reviews of all the tasks, scores and, above all, comments. It means you can learn from the thoughts of other classmates." (C110)

Regarding ideas to improve feedback, there is a frequency of 5 or over for comments that the reports are processed and displayed very slowly (8), and that the information for an assessment is available for everyone in this work group (5). In this respect, it should be emphasised that this is an option, configured by the teacher when implementing the environment. Regarding the first question, one participant in the focus group mentioned the following:

> "As a personal example, [the report] took a long time to load, when I tried to view it, it took a pretty long time to load and often ended up not working properly. So, I have looked at it a few times but not too many." (GF3).

This perception of the need to improve the reports relates to the dimension in the questionnaire referring to satisfaction and interest regarding the reports generated by EvalCOMIX-FLOASS (DF4), that received a positive perception of 74.9%, a little under the general perception of the other items.

In terms of evaluative judgement (EJ), high satisfaction is observed among the students, with more than 84% acceptance in three out of the four dimensions (EJ2, EJ3 and EJ4). In qualitative terms, although a total of 11 different improvement proposals were compiled, they all boiled down to 1, 2 or 3 references. However, students greatly appreciated the positive aspects in this field. A total of 29 refer explicitly to the fact that this is a tool which encourages student participation in the assessment, and also that it is very useful to find out about the assessment instruments with their corresponding criteria and assessment aspects (21) and that this directly benefits the evaluating skills of the students that use it (14). Examples of these answers are given below:

> "The tool reveals which assessment criteria are used and how they are applied to your task, more clearly and simply than usual criteria. It also makes you part of your own assessment process, encouraging you to become familiar with your own mistakes and aspects to be improved in future work" (C111).

General satisfaction with the environment is also positive, with 82.2% of students satisfied or very satisfied with the EvalCOMIX-FLOASS tools to evaluate their tasks. Qualitatively, it should be highlighted that 21 students considered that no aspect required improvement on the EvalCOMIX-FLOASS environment.

Finally, it is important to mention that the students would like the EvalCOMIX-FLOASS environment to be used in other university subjects, as deduced from the median of 5 obtained in the Generalisation dimension, and that 77.5% of the students agreed or agreed completely with this statement. In this field, none of the improvements or the qualitative positive aspects in this regard obtained at least 5 references.

4.2.2 Differences in the Students' Perceptions

The students' responses to the questionnaire are analysed below, according to their level, gender and age, using a PERMANOVA test. The results in Table 6 show significant differences in the students' satisfaction with the environment overall, in the satisfaction and in the generalisation, depending on whether they are taking a degree or a master's degree ($p < .05$). Specifically, the perception from master's degree students is more positive than from the degree students. In the same way, significant differences are seen regarding all dimensions depending on gender ($p < .05$). In this case, the perception of female students is more positive than male students. Finally, no significant differences were seen when analysing age-dependent differences.

To dig deeper into the analysis of the differences encountered, the Kruskal–Wallis test was run on each of the questionnaire items, and the results are presented in Table 7.

Significant differences are confirmed in various items depending on the students' level and gender although no age-dependent differences were observed. In detail, significant differences were seen in four items by level (FE4, EJ3, SA1 and GE1), obtaining a more positive assessment from the master's degree students than degree students. While the average score from the degree students for these four dimensions was 4.37, 3.92, 4.12 and 4.02 respectively, the average score from the master's degree students was 4.64, 4.34, 4.60 and 4.56 respectively. On the other hand, significant differences were observed in five items, depending on gender (FE4, EJ3, EJ4, SA2 and GE1), where female students gave a more positive assessment than the male students. While the average score from the female students was 4.52, 4.11, 4.30, 4.50 and 4.27 respectively, for the male students it was 4.18, 3.71, 3.96, 4.16 and 3.75 respectively. It stands out that three of these items match for

Table 6 PERMANOVA results depending on level, gender and age of the students

Dimensions		Level			Gender			Age		
		F	Sig	R^2	F	Sig	R^2	F	Sig	R^2
Overall dimensions	P-I	4.501	.016	.016	9.40	.001	.033	1.32	.237	.009
DF Items	P-II	0.061	.540	.002	4.57	.022	.016	1.37	.222	.009
FE Items		2.53	.065	.009	5.21	.004	.019	0.34	.921	.002
EJ Items		1.566	.183	.005	5.37	.005	.020	0.43	.837	.003
SA Items		10.787	.003	.038	10.83	.001	.038	0.47	.674	.003
GE Items		8.226	.005	.029	11.26	.001	.040	2.48	.094	.018
CATEF_E Items	P-III	2.776	.046	.010	7.05	.001	.025	0.82	.513	.006

Table 7 Results from the Kruskall-Wallis test depending on level, gender and age of the students

	Level			Gender			Age		
	H	Sig	η_H^2	H	Sig	η_H^2	H	Sig	η_H^2
Design-Facility of use (DF)									
DF1	.525	.469	.002	2.523	.112	.009	2.164	.339	.008
DF2	.145	.704	.001	1.267	.260	.005	1.662	.436	.006
DF3	.166	.683	.001	2.925	.087	.011	1.371	.504	.005
Feedback (FE)									
FE1	.634	.426	.002	2.880	.090	.011	.910	.635	.000
FE2	.168	.682	.001	.495	.482	.002	2.122	.346	.004
FE3	.130	.719	.000	1.227	.268	.004	1.269	.530	.008
FE4	7.145	.008	.026	6.898	.009	.025	.452	.798	.011
Evaluative judgement (EJ)									
EJ1	.227	.634	.001	1.574	.210	.006	.098	.952	.003
EJ2	3.151	.076	.012	.039	.843	.000	1.185	.553	.008
EJ3	6.062	.014	.022	8.549	.003	.031	2.209	.331	.005
EJ4	.318	.573	.001	4.314	.038	.016	3.107	.212	.002
Satisfaction (SA)									
SA1	9.576	.002	.035	2.783	.095	.010	.164	.921	.001
SA2	1.377	.241	.005	9.195	.002	.034	.566	.754	.002
Generalisation (GE)									
GE1	7.857	.005	.029	6.080	.014	.022	2.356	.308	.009

differences in level and gender, specifically, the interest and satisfaction in the tool's reports (FE4), the tool's capacity to encourage participation (EJ3) and the interest in generalising the use of the tool in other subjects (GE1).

4.2.3 The Teachers' Perception

The results were extremely positive when questioning teachers on their use of EvalCOMIX-FLOASS (see Table 8). Average scores of 4.67 and 4.64 were obtained in the feedback (FE) and evaluative judgement (EJ) dimensions, respectively, which indicates a high degree of satisfaction. Out of the 18 items considered in the questionnaire, 14 obtained a median of 5, which demonstrates the usefulness of the tool to efficiently evaluate whether the learning outcomes and skill development have been achieved, and to what extent, in addition to producing valuable reports on improving learning and assessment.

These results are backed up by 100% of the teachers, who stated they agree or agree completely on three key dimensions. Firstly, DF1, related to the facility of use (DF), indicates that EvalCOMIX can be used to design assessment instruments. Secondly, FE1, related to feedback, highlights that the environment makes it easier to monitor and systematically assess the learning outcomes and skill development Finally, EJ1, within evaluative judgement, confirms that EvalCOMIX facilitates assessment instruments to evaluate the level of achievement for learning outcomes and skill development.

However, two areas for improvement were identified by the teachers: the possibility of using EvalCOMIX with minimal instructions (79.1%) and making instrument design more intuitive (76.8%). However, all the other dimensions obtained scores over 88%.

Table 8 Central trend and dispersion measurements in the dimensions and items of the CATEF_P questionnaire

Dimensions and items	Mdn	M	SD
Design-Facility of use (DF)	*4.50*	*4.40*	*.58*
DF1 I can design assessment instruments using EvalCOMIX-FLOASS	5.00	4.86	.35
DF2 I can use the EvalCOMIX-FLOASS tool with minimal instructions	4.00	4.23	.84
DF3 It is intuitive to design assessment instruments using EvalCOMIX-FLOASS	4.00	4.14	.91
DF4 It is easy to design assessment instruments using EvalCOMIX-FLOASS	5.00	4.35	.78
Feedback (FE)	*5.00*	*4.67*	*.41*
FE1 EvalCOMIX-FLOASS allows me to monitor and assess the learning outcomes and skill development in an organised and systematic way	5.00	4.81	.39
FE2 The reports generated using information stored in EvalCOMIX-FLOASS are interesting and I am satisfied with them	5.00	4.53	.63
FE3 I am satisfied that the information provided by EvalCOMIX-FLOASS can be used to improve the learning and the assessment	5.00	4.65	.57
Evaluative judgement (EJ)	*5.00*	*4.64*	*.48*
EJ1 EvalCOMIX-FLOASS allows me to use assessment instruments to assess the level of achievement of the learning outcomes and skill development	5.00	4.74	.44
EJ2 EvalCOMIX-FLOASS makes me more efficient when obtaining evidence on how far the learning outcomes and skill development have been achieved	5.00	4.58	.70
EJ3 With EvalCOMIX-FLOASS, it is easy to assess the level of achievement for learning outcomes and skill development	5.00	4.60	.62
Satisfaction (SA)	*5.00*	*4.54*	*.56*
SA1 I am satisfied with the EvalCOMIX-FLOASS application to design assessment instruments	5.00	4.51	.67
SA2 I am satisfied with the EvalCOMIX-FLOASS application to evaluate the level of achievement of the learning outcomes and skill development	5.00	4.58	.66
SA3 I am satisfied with the usefulness, ease and interest of EvalCOMIX-FLOASS for the university community	5.00	4.49	.74
Generalisation (GE)	*4.60*	*4.43*	*.57*
GE1 Teachers	5.00	4.67	.61
GE2 Subject coordination	5.00	4.60	.69
GE3 Degree coordination	5.00	4.53	.80
GE4 Deanships and head teachers	4.00	4.19	.79
GE5 Assessment and accreditation agencies	4.00	4.16	.87

In terms of generalised EvalCOMIX use, the teachers agreed or agreed completely, 90.7% and 93.1% respectively, that it should be implemented by teachers, subject coordinators and degree coordinators. In contrast, only 74.4% thought that accreditation agencies should use the tool.

4.2.4 Differences in the Teachers' Perceptions

The PERMANOVA test was performed to pinpoint significant differences in teachers' perception regarding the various dimensions of the EvalCOMIX-FLOASS environment (Table 9), depending on the university, gender and age. In general, the results indicate that there are barely any differences among the questionnaire dimensions. It can only be highlighted that significant differences were found in the dimension related to the evaluative judgement (EJ) depending on gender ($H = 4.670$, $p < .05$).

Kruskall-Wallis test was performed to evaluate the possible differences in the items considered individually regarding the university, gender and age (Table 10). In the evaluative judgement dimension, significant differences were found in the teachers' perception relating to their gender in item EJ2, which indicates that the tool helped them to become more efficient when obtaining evidence of the degree of achieving learning outcomes and skills development for their students ($H = 11.352$, $p < .05$). Specifically, the average score from female teachers is significantly higher than from male teachers (4.84 and 4.37 respectively). Regarding the university, relatively high values of H and ηH2 were seen in the FE2 (12.369, $p > .05$) and GE2 ($H = 11.506$, $p < .05$) items, which suggests the possibility of substantial differences between the groups. However, the level of significance is slightly higher than 0.05 in both cases, so we cannot reject the null hypothesis that there is no relationship between the university and the response given.

4.2.5 Differences in Perception Between Students and Teachers

Table 11 shows Kruskal–Wallis test results comparing student and teacher perceptions in five dimensions. The students perceive the DF dimension more positively than the teachers, while the teachers have a more positive perception of the FE, EJ and SA dimensions. In the GE dimension, students and teachers demonstrate similar perceptions, with no significant differences.

4.3 The Mediating Nature of Feedback and Evaluative Judgement

Regarding the relationship between satisfaction with the feedback and evaluative judgement, Fig. 7 presents the results for the analysis of the students' perception. This demonstrates that 73.5% of the variance (R^2) in the satisfaction construct (SA) is explained by the DF, FE and EJ constructs. This model presents an SRMR of 0.056, which indicates an appropriate level of adjustment as it is below the threshold value of 0.08. The *path* values were significant in all cases ($p < .05$).

This is confirmed as a multiple mediation where the feedback (FE) and the evaluative judgement (EJ) operate as mediating variables between design-facility (DF) and satisfaction (SA). As shown in Table 12, DF has a direct significant effect (0.316) ($t = 17.174$, $p < .05$) on SA and the total indirect effect of the relationships between DF and SA (0.434) is also significant ($t = 9.217$, $p < .05$). The analysis of the indirect effects shows us how the feedback and the evaluative judgement intervene in the mediation and how mediating

Table 9 PERMANOVA results depending on the university, gender and age of the teachers

Dimensions		University			Gender			Age		
		F	Sig	R^2	F	Sig	R^2	F	Sig	R^2
Overall	P-I	2.0711	.042	.025	1.7448	.129	.080	1.2641	.247	.117
DF Items	P-II	1.3198	.232	.180	1.4834	.184	.069	1.4332	.197	.131
FE Items		1.8396	.060	.235	0.4701	.839	.023	1.0073	.423	.096
EJ Items		1.0203	.437	.145	4.670	.011	.189	0.6717	.722	.066
SA Items		1.8371	.073	.234	0.2503	.945	.012	1.2188	.287	.114
GE Items		1.7999	.050	.231	1.8315	.101	.084	0.8097	.604	.079
CATEF_E Items	P-III	0.9267	.421	.022	1.619	.111	.075	1.0401	.370	.099

Table 10 Results from the Kruskall-Wallis test depending on university, gender and age of the teachers

	University			Gender			Age		
	H	Sig	η_H^2	H	Sig	η_H^2	H	Sig	η_H^2
Design-Facility of use (DF)									
DF1	10.002	.125	.238	3.134	.209	.075	3.596	.463	.086
DF2	7.878	.247	.188	3.241	.198	.077	5.496	.240	.131
DF3	10.596	.102	.252	3.325	.190	.079	6.569	.160	.156
DF4	2.500	.868	.060	3.525	.172	.084	7.723	.102	.184
Feedback (FE)									
FE1	4.160	.655	.099	4.735	.094	.113	2.337	.674	.056
FE2	12.369	.054	.294	.044	.978	.001	5.376	.251	.128
FE3	8.009	.237	.191	.354	.838	.008	3.561	.469	.085
Evaluative judgement (EJ)									
EJ1	3.862	.695	.092	1.275	.529	.030	2.365	.669	.056
EJ2	3.896	.691	.093	11.352	.003	.270	4.279	.370	.102
EJ3	7.493	.278	.178	1.218	.544	.029	2.636	.620	.063
Satisfaction (SA)									
SA1	9.657	.140	.230	.392	.822	.009	5.168	.271	.123
SA2	10.521	.104	.250	.260	.878	.006	4.459	.347	.106
SA3	9.674	.139	.230	.672	.714	.016	6.504	.165	.155
Generalisation (GE)									
GE1	8.657	.194	.206	4.951	.084	.118	2.809	.590	.067
GE2	11.506	.074	.274	1.199	.549	.029	6.644	.156	.158
GE3	8.112	.230	.193	2.720	.257	.065	6.490	.165	.155
GE4	8.168	.226	.194	2.639	.267	.063	1.317	.859	.031
GE5	6.535	.366	.156	3.068	.217	.073	.571	.966	.014

variables are consistent with the hypothesised model, although as Zhao et al. (2010) state, there might be other mediators, not considered here, that complete this model.

The results for the teachers are presented in Fig. 8, showing that 65.9% of the variance (R^2) of the satisfaction construct (SA) is explained by the DF, FE and EJ constructs.

However, this model presents an SRMR of 0.09 and due to the small sample size (n = 43), it was not possible to calculate the inverse matrix as there are fewer observations than degrees of freedom, meaning that the momentum matrix cannot be inverted. However, similar results are obtained for the direct and indirect effects as seen for the students.

5 Discussion and Future Direction Research

This study was intended to validate the EvalCOMIX-FLOASS technological environment, by answering two research questions. Firstly, it considered the degree of technological acceptance among students and teachers and the possible differences in this degree of acceptance. The questionnaire results, from both closed and qualitative questions, and the focus groups, demonstrate a high degree of acceptance among the students, who clearly and notably perceive the benefits provided by the EvalCOMIX-FLOASS environment when developing their evaluative judgement and in the feedback received and provided. Slight differences of opinion have been found, such as more favourable opinions among master's degree students compared to undergraduate degree students, likewise among women compared to men. In turn, the teachers demonstrate a high degree of acceptance concerning the environment, and their opinions barely differ.

The second question enquired about a possible relationship between user satisfaction with the EvalCOMIX-FLOASS environment and evaluative judgement and feedback. In this respect, both evaluative judgement and feedback clearly operate as mediating variables between the design and satisfaction with use of the environment.

While use of the technology was emphasised to achieve an efficient and consistent assessment, little attention has been paid to using the technology to rethink relationships between learning and assessment and the role that technology might play to support these changes (Timmis et al., 2016). Changes or innovations, as raised by these authors, suggest the need to develop new tools that reflect the main pedagogic innovators, technologies that support more democratic, fair and unbiased assessment methods. The challenge lies in incorporating the potential nature of the technologies that improve the assessment (TEAs) and the LA to provide students with new decision-making opportunities, inclusion of participatory evaluation methods such as self-assessment or peer assessment, with the aim of developing their evaluative judgement (Boud et al., 2018a, 2018b) and thereby strengthening self-assessment and lifelong learning.

Table 11 Descriptors and outcomes from the Kruskal–Wallis test when comparing students and teachers' perceptions

Dimensions	Students			Teachers			K-W		
	Mdn	M	SD	Mdn	M	SD	H	Sig	η_H^2
DF	4.33	4.13	.82	3.60	3.52	.46	36.466	.001	.067
FE	4.25	4.18	.69	5.00	4.67	.41	10.120	.001	.027
EJ	4.32	4.13	.68	5.00	4.64	.48	24.411	.001	.060
SA	4.25	4.14	.84	5.00	4.54	.56	12.199	.001	.028
GE	5.00	4.12	1.20	4.60	4.43	.57	.126	.722	.009

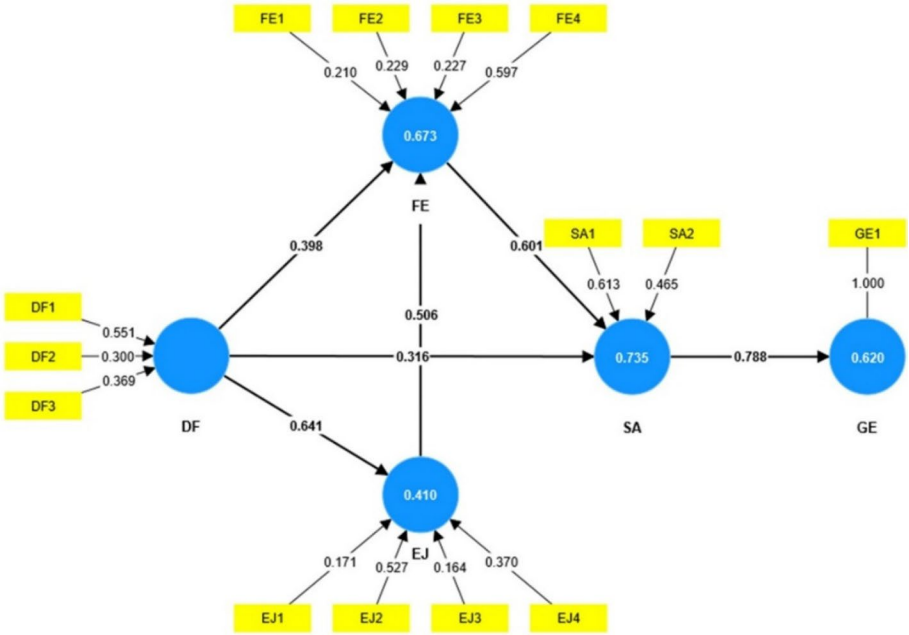


Fig. 7 Results from the structural model of the students' perception

Curricular activities must be structured to offer students continuous opportunities to obtain, process and apply feedback (Bikanga Ada, 2023; Malecka et al., 2022; Van der Kleij & Lipnevich, 2021). The assessment-as-learning focus, which encompasses the Eval-COMIX-FLOASS environment, not only facilitates continuous and significant feedback, but teachers can also provide personalised and timely feedback to the students, helping them to understand their progress and address areas that require greater attention.

Furthermore, they provide technology that helps students to monitor their progress and explore their strengths and weaknesses; it empowers them by adapting their learning process to individual needs (Tsai et al., 2020a). The evidence compiled in the use of Eval-COMIX-FLOASS shows that not only were they able to monitor their own learning, but they also state that this gave them the chance to think about and assess their own learning.

Table 12 Summary of confirming the effect of mediation of DF on SA

	Effect	95% CI	*t*	*p*	Type of mediation
Total effect					
DF → SA	0.750	[0.344, 0.529]	17.174	0.000	
Direct effect					
DF → SA	0.316	[0.199, 0.419]	5.681	0.000	
Total indirect effect					
DF → SA	0.434	[0.344, 0.529]	9.217	0.000	
Specific indirect effect					
DF → FE → SA	0.239	[0.023, 0.072]	3.662	0.000	Complementary
DF → EJ → FE → SA	0.195	[0.277, 0.393]	11.378	0.000	Complementary

Therefore, the EvalCOMIX-FLOASS environment is not only contributing to their academic development, but it also promotes their evaluative judgement, helping them become aware of their learning process and make better-informed decisions on how to improve. In this respect, we should insist that students' evaluative judgement can only be developed through an in-depth change in how assessment and feedback are usually conceptualised and implemented in our universities (Panadero et al., 2019a).

Based on these findings, we recommend that teachers actively integrate feedback mechanisms into their course design, ensuring that students receive timely and constructive feedback aligned with learning outcomes. Additionally, students should be encouraged to use EvalCOMIX-FLOASS to track their progress and reflect on their evaluative judgements, fostering a deeper understanding of their learning processes.

Technologically, it is a challenge to integrate multiple data origins in class group management, without diminishing the application of LA techniques (Samuelsen et al., 2019). As shown in the interviews and the usability questionnaires, implementing the EvalCOMIX-FLOASS environment as a plugin for LMS Moodle has meant that teachers and students do not need to move between different platforms, respecting the centrality of the LMS provided by the academic institution and making academic life easier for both.

The results presented must be considered with some limitations. Firstly, most of the teachers come from a Social Sciences background and, among the students, there are many more undergraduate degree students than master's degree students (81.8% vs 18.2%) and an over-representation of women compared to men (70.9% vs 29.1%). These sample characteristics limit the chances of generalising the results.

A second aspect to weigh up is the environment design, as students made proposals for aesthetic and functional improvements. In fact, some of the functional features they proposed have already been set up and can be implemented. Consequently, despite the high intensity of the environment use, activities should be developed which give teachers and students optimum use of the environment's full potential. In this respect, a new version

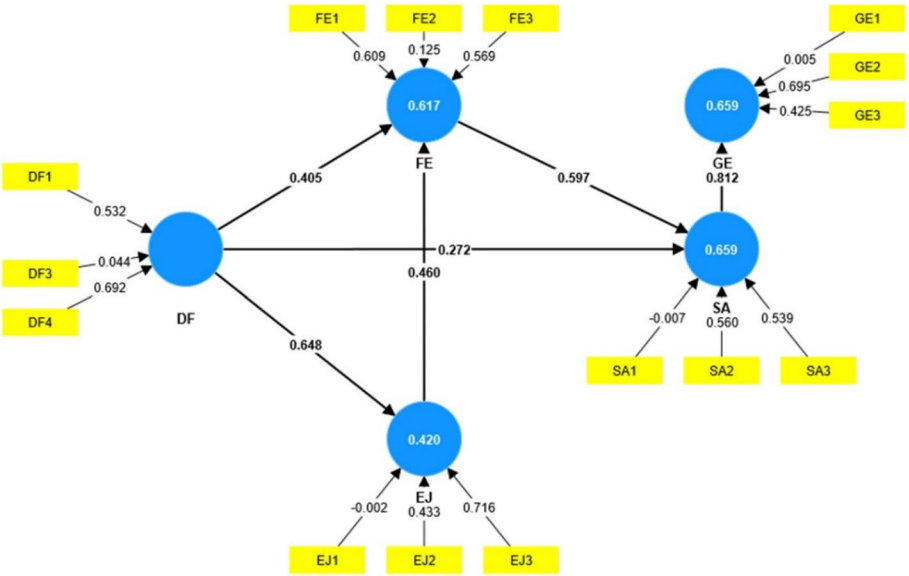


Fig. 8 Results from the structural model of the teachers' perception

currently in development takes the proposals into consideration and it will be important to check the impact of these improvements on users' perceptions.

Finally, with a view to future research, we should look in greater depth at the impact of artificial intelligence on assessment processes and how this might relate to the EvalCOMIX-FLOASS environment. Current AI-based assessment systems are often isolated and designed for specific tasks. However, future systems could harness Multi-Agent Systems (MAS) to deliver adaptive, context-aware evaluations that consider both the learner's profile and the specific context of a subject, degree programme or training area (Viswanathan et al., 2022). This approach would help address pedagogical and ethical challenges by ensuring that AI-generated feedback enhances, rather than replaces, human feedback. It is relevant to explore the repercussions, possibilities and limitations in this field (Bearman et al., 2024; García-Peñalvo, 2023; Zhai, 2023). More specifically, it will be necessary to examine how EvalCOMIX-FLOASS differs from or complements AI tools in its ability to foster evaluative judgement and supply high-quality feedback. It is also important to research perceptions of AI-based evaluation systems among students and teachers and analyse how these perceptions influence the acceptance and use of these systems. Careful consideration must be given to pedagogical challenges, such as the need for teacher training and the importance of preserving human interaction in the learning process, and to ethical challenges, such as data privacy and the risk of algorithmic bias.

6 Conclusion

This study presented the evaluation of the EvalCOMIX-FLOASS technological environment which was developed to facilitate monitoring and assessment of learning outcomes. It has been shown that this technology facilitates development of evaluative judgement and feedback processes, both essential characteristics of the assessment-as-learning focus. Consequently, teachers are offered a useful environment which helps them to monitor and assess the level of achievement from the learning outcomes. Furthermore, it allows students to assess their learning, the strategies that they use to guide this learning process, and the mechanisms required to move forwards in their own learning. In short, it is a technological environment which facilitates lifelong learning.

Despite these benefits, it is important to recognise the study limitations. Although diverse in terms of universities and disciplines, the sample mainly focuses on the Social Sciences area, which might influence the generalisation of the results. Furthermore, use of the EvalCOMIX-FLOASS environment requires greater analysis in future studies to assess its applicability in other educational contexts and disciplines.

Finally, with a view to future research projects, we should look in greater depth at the impact of artificial intelligence on evaluation processes and how this might relate to the EvalCOMIX-FLOASS environment. It is relevant to explore the repercussions, possibilities and limitations in this field to adapt technological tools better to the current and future educational needs.

Author Contributions All authors contributed to the study conception and design. Material preparation and data collection performed by María Soledad Ibarra-Sáiz and Gregorio Rodríguez-Gómez. Qualitative data analysis performed by Miguel Ángel Gómez-Ruiz and quantitative data analysis performed by Antonio

Balderas and Gregorio Rodríguez-Gómez. The first draft of the manuscript was written by María Soledad Ibarra-Sáiz and Gregorio Rodríguez-Gómez. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Funding for open access publishing: Universidad de Cádiz/CBUA. This work was made possible by the FLOASS Project—*Learning Outcomes and Learning Analytics in Higher Education: An Action Framework from Sustainable Assessment(Ref. RTI2018-093630-B-I00) funded by the Spanish Ministry of Science, Innovation and Universities (MCIU), the State Research Agency (AEI), the European Regional Development Fund (ERDF); and by the ValeFLOASS Project—Valorisation of the e-FLOASS environment for the monitoring and assessment of learning outcomes in the educational system(Ref. PDC2022-133050-I00) funded by the Spanish Ministry of Science and Innovation (MCIN), the State Research Agency (AEI/*https://doi.org/10.13039/501100011033) and the European Union "NextGeneration EU"/PRTR. These projects have been led by the UNESCO Chair in Assessment & Evaluation, Innovation and Excellence in Education from the University of Cadiz.

Data Availability The data supporting this study's finding are available from the corresponding author upon reasonable request.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Afzaal, M., Zia, A., Nouri, J., & Fors, U. (2024). Informative feedback and explainable AI-based recommendations to support students' self-regulation. *Technology, Knowledge and Learning, 29*(1), 331–354. https://doi.org/10.1007/s10758-023-09650-0

Ajjawi, R., & Boud, D. (2018). Examining the nature and effects of feedback dialogue. *Assessment & Evaluation in Higher Education, 43*(7), 1106–1119. https://doi.org/10.1080/02602938.2018.1434128

Anderson, M. J. Permutational multivariate analysis of variance (PERMANOVA). Wiley StatsRef: Statistics Reference Online, 1–15 (2017). https://doi.org/10.1002/9781118445112.stat07841

Balderas, A., Palomo-Duarte, M., Dodero, J. M., Ibarra-Sáiz, M. S., & Rodríguez-Gómez, G. (2018). Scalable authentic assessment of collaborative work assignments in wikis. *International Journal of Educational Technology in Higher Education*, 15, 40. https://doi.org/10.1186/s41239-018-0122-1

Bearman, M., Boud, D., & Ajjawi, R. (2020). New directions for assessment in a digital world. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining university assessment in a digital world* (pp. 7–18). Springer. https://doi.org/10.1007/978-3-030-41956-1_2

Bearman, M., & Luckin, R. (2020). Preparing university assessment for a world with AI: Tasks for human intelligence. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining university assessment in a digital world* (pp. 79–94). Springer. https://doi.org/10.1007/978-3-030-41956-1_5

Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment and Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2024.2335321

Berk, A., Zhang, K., Gizem, F., & Yilmaz, R. (2022). Learning analytics based feedback and recommendations in flipped classrooms: An experimental study in higher education. *Journal of Research on Technology in Education*. https://doi.org/10.1080/15391523.2022.2040401

Bikanga Ada, M. (2023). Evaluation of a mobile web application for assessment feedback. *Technology, Knowledge and Learning, 28*, 23–46. https://doi.org/10.1007/s10758-021-09575-6

Boud, D., Ajjawi, R., Dawson, P., & Tai, J. (Eds.). (2018a). *Developing evaluative judgement in higher education*. Routledge.

Boud, D., Dawson, P., Tai, J., & Ajjawi, R. (2018b). Creating an agenda for developing students' evaluative judgement. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement: Assessment for knowing and producing quality work* (pp. 186–195). Routledge. https://doi.org/10.4324/9781315109251-20

Boud, D., & Molloy, E. (2013). Re-thinking models of feedback for learning: The challenge of design. *Assessment and Evaluation in Higher Education, 38*(6), 698–712. https://doi.org/10.1080/02602938.2012.691462

Cain, A., Tubino, L., & Krishnan, S. (2020). Using technology to enable a shift from marks to outcomes-based assessment. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining university assessment in a digital world* (pp. 229–245). Springer. https://doi.org/10.1007/978-3-030-41956-1_16

Caird, S., & Lane, A. (2015). Conceptualising the role of information and communication technologies in the design of higher education teaching models used in the UK. *British Journal of Educational Technology, 46*(1), 58–70.

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education, 36*(4), 395–407. https://doi.org/10.1080/03075071003642449

Chaparro-Peláez, J., Iglesias-Pradas, S., Rodríguez-Sedano, F. J., & Acquila-Natale, E. (2019). Extraction, processing and visualization of peer assessment data in moodle. *Applied Sciences, 10*(1), 163. https://doi.org/10.3390/app10010163

Creswell, J. W. (2015). *A concise introduction to mixed methods research*. SAGE Publications.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Decuir-Gunby, J. T., & Schutz, P. A. (2017). *Developing a mixed methods proposal*. SAGE Publications.

Deepak, K. C. (2017). Evaluation of moodle features at kajaani university of applied sciences–case study. *Procedia Computer Science, 116*, 121–128. https://doi.org/10.1016/j.procs.2017.10.021

Deeva, G., Bogdanova, D., Serral, E., Snoeck, M., & De Weerdt, J. (2021). A review of automated feedback systems for learners: Classification framework, challenges and opportunities. *Computers and Education, 162*(20), 104094. https://doi.org/10.1016/j.compedu.2020.104094

Diamantopoulos, A. (2006). The error term in formative measurement models: Interpretation and modeling implications. *Journal of Modelling in Management, 1*(1), 7–17. https://doi.org/10.1108/17465660610667775

Farrell, T., & Rushby, N. (2016). Assessment and learning technologies: An overview. *British Journal of Educational Technology, 47*(1), 106–120.

Flick, U. (2014). Mapping the field. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 3–18). SAGE Publications.

García-Peñalvo, F. J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT: disrupción o pánico. *Education in the Knowledge Society (EKS), 24*, e31279. https://doi.org/10.14201/eks.31279

Gómez-Ruiz, M.A., & Quesada-Serra, V. (2020). An Analysis of Shared Grading in Co-Assessment Practices by Teachers and Students. *RELIEVE, 26* (1), art. M6. https://doi.org/10.7203/relieve.26.1.16567

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS_SEM)*. SAGE Publications.

Hassenzahl, M. User experience (UX): Towards an experiential perspective on product quality. In *Proceedings of the 20th international conference of the association Francophone d'Interaction Homme-Machine* (2008). https://doi.org/10.1145/1512714.1512717

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Henderson, M., Ajjawi, R., Boud, D., & Molloy, E. (Eds.). (2019). *The impact of feedback in higher education*. Springer. https://doi.org/10.1007/978-3-030-25112-3

Henseler, J. (2021). *Composite-based structural equation modeling. Analyzing latent and emergent variables*. Guilford Press.

Ibarra-Sáiz, M.S., & Rodríguez-Gómez, G. (2019). Evaluación como aprendizaje. In J.Paricio, A. Fernández & I. Fernández (Eds.), *Cartografía de la buena docencia universitaria. Un marco para el desarrollo del profesorado basado en la investigación* (pp. 175-196). Narcea

Ibarra-Sáiz, M. S. & Rodríguez-Gómez, G. (2020). EvalCOMIX®: A Web-Based Programme to Support Collaboration in Assessment. In I. Management Association (Ed.), *Learning and Performance Assessment: Concepts, Methodologies, Tools, and Applications* (pp. 455-481). IGI Global. https://doi.org/10.4018/978-1-7998-0420-8.ch022

International Organization for Standardization. (2019). *ISO 9241-210:2019. Ergonomics of human-system interaction—Part 210: Human-centred design for interactive systems*.

Jensen, L. X., Bearman, M., & Boud, D. (2021). Understanding feedback in online learning—A critical review and metaphor analysis. *Computers & Education, 173*, Article 104271. https://doi.org/10.1016/j.compedu.2021.104271

Johnson, R. L., & Morgan, G. B. (2016). *Survey scales. A guide to development, analysis, and reporting*. The Guilford Press.

Kärner, T., Warwas, J., & Schumann, S. A. (2021). Learning analytics approach to address heterogeneity in the classroom: The teachers' diagnostic support system. *Technology, Knowledge and Learning, 26*, 31–52. https://doi.org/10.1007/s10758-020-09448-4

Khosravi, H., Gyamfi, G., Hanna, B. E., Lodge, J., & Abdi, S. (2021). Bridging the gap between theory and empirical research in evaluative judgment. *Journal of Learning Analytics, 8*(3), 117–132. https://doi.org/10.18608/jla.2021.7206

Kobayashi, M. (2020). Does anonymity matter? Examining quality of online peer assessment and students' attitudes. *Australasian Journal of Educational Technology, 36*(1), 98–110. https://doi.org/10.14742/ajet.4694

López-Pastor, V., & Sicilia-Camacho, A. (2017). Formative and shared assessment in higher education. Lessons learned and challenges for the future. *Assessment & Evaluation in Higher Education, 42*(1), 77–97. https://doi.org/10.1080/02602938.2015.1083535

Malecka, B., Boud, D., & Carless, D. (2022). Eliciting, processing and enacting feedback: Mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education, 27*(7), 908–922. https://doi.org/10.1080/13562517.2020.1754784

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Sage Publications.

Morrell, L. J. (2021). Iterated assessment and feedback improves student outcomes. *Studies in Higher Education, 46*(3), 485–496. https://doi.org/10.1080/03075079.2019.1643301

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education, 39*(1), 102–122. https://doi.org/10.1080/02602938.2013.795518

Panadero, E., Broadbent, J., Boud, D., & Lodge, J. M. (2019a). Using formative assessment to influence self- and co-regulated learning: The role of evaluative judgement. *European Journal of Psychology of Education, 34*, 535–557. https://doi.org/10.1007/s10212-018-0407-8

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129–144. https://doi.org/10.1016/j.edurev.2013.01.002

Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review, 35*, Article 100416. https://doi.org/10.1016/j.edurev.2021.100416

Panadero, E., Lipnevich, A., & Broadbent, J. (2019b). Turning self-assessment into self-feedback. In M. Henderson, R. Ajjawi, D. Boud, & E. Molloy (Eds.), *The impact of feedback in higher education* (pp. 147–164). Springer. https://doi.org/10.1007/978-3-030-25112-3_9

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology, 50*(1), 128–138. https://doi.org/10.1111/bjet.12592

Peña-Ayala, A. (Ed.). (2017). *Learning analytics: Fundaments, applications, and trends. A view of the current state of the art to enhance e-learning*. Springer.

Petersen, F. & Groenewald, B. (2021). Students' engagement in anonymous peer review: Using the open-source Sakai platform. https://doi.org/10.48550/arXiv.2108.09955

Pitt, E., & Winstone, N. (2020). Towards technology enhanced dialogic feedback. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining university assessment in a digital world* (pp. 79–94). Springer. https://doi.org/10.1007/978-3-030-41956-1_7

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Ringle, C. M., Wende, S., & Becker, J.M. (2022). *SmartPLS 4*. SmartPLS GmbH. http://www.smartpls.com

Rodríguez-Gómez, G., Gil-Flores, J., & García-Jiménez, E. (1999). *Metodología de la investigación cualitativa* (3rd ed.). Aljibe.

Rodríguez-Gómez, G., Gómez-Ruiz, M.A. e Ibarra-Sáiz, M.S. (2022, 18 octubre). *Entorno e-FLOASS para el diseño, seguimiento y evaluación de los resultados de aprendizaje* [Vídeo]. YouTube. https://youtu.be/3P-lJLbA62E

Rodríguez, M. F., Nussbaum, M., Yunis, L., Reyes, T., Alvares, D., Joublan, J., & Navarrete, P. (2022). Using scaffolded feedforward and peer feedback to improve problem-based learning in large classes. *Computers and Education, 182*, Article 104446. https://doi.org/10.1016/j.compedu.2022.104446

Rönn, C., & Pettersson, D. (2023). Swedish students' everyday school life and teachers' assessment dilemmas: Peer strategies for ameliorating schoolwork for assessment. *Educational Assessment, Evaluation and Accountability, 35*, 37–66. https://doi.org/10.1007/s11092-022-09400-3

Ruelmann, M., Charalambous, C. Y., & Praetorius, A. K. (2023). The representation of feedback literature in classroom observation frameworks: An exploratory study. *Educational Assessment, Evaluation and Accountability, 35*, 67–104. https://doi.org/10.1007/s11092-022-09403-0

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education, 35*(5), 535–550. https://doi.org/10.1080/02602930903541015

Salazar, G., & Russi-Vigoya, M. N. (2021). Technology Readiness Level as the Foundation of Human Readiness Level. *Ergonomics in Design, 29*(4), 25–29. https://doi.org/10.1177/10648046211020527

Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—Review of literature. *Research and Practice in Technology Enhanced Learning, 14*, 11. https://doi.org/10.1186/s41039-019-0105-4

Smith, C. D., Worsfold, K., Davies, L., Fisher, R., & McPhail, R. (2013). Assessment literacy and student learning: The case for explicitly developing students 'assessment literacy.' *Assessment & Evaluation in Higher Education, 38*(1), 44–60. https://doi.org/10.1080/02602938.2011.598636

Sullivan, D., & Watson, S. (2015). Peer assessment within hybrid and online courses: Students' view of its potential and performance. *Journal of Educational Issues*. https://doi.org/10.5296/jei.v1i1.7255

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education, 76*, 467–481. https://doi.org/10.1007/s10734-017-0220-3

Tan, J. S., & Chen, W. (2022). Peer feedback to support collaborative knowledge improvement: What kind of feedback feed-forward? *Computers and Education, 187*, Article 104467. https://doi.org/10.1016/j.compedu.2022.104467

Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal, 42*(3), 454–476.

Tsai, Y. S., Perrotta, C., & Gašević, D. (2020a). Empowering learners with personalised learning approaches? Agency, equity and transparency in the context of learning analytics. *Assessment & Evaluation in Higher Education, 45*(4), 554–567. https://doi.org/10.1080/02602938.2019.1676396

Tsai, Y. S., Rates, D., Moreno-Marcos, P. M., Muñoz-Merino, P. J., Jivet, I., Scheffel, M., Drachsler, H., Delgado Kloos, C., & Gašević, D. (2020b). Learning analytics in European higher education—Trends and barriers. *Computers & Education, 155*, Article 103933. https://doi.org/10.1016/J.COMPEDU.2020.103933

Van der Kleij, F. M., & Lipnevich, A. A. (2021). Student perceptions of assessment feedback: A critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability, 33*, 345–373. https://doi.org/10.1007/s11092-020-09331-x

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Verdú, M. J., De Castro, J. P., Regueras, L. M., & Corell, A. (2021). MSocial: Practical integration of social learning analytics into Moodle. *IEEE Access, 9*, 23705–23716. https://doi.org/10.1109/ACCESS.2021.3056914

Viswanathan, N., Meacham, S., & Adedoyin, F. F. (2022). Enhancement of online education system by using a multi-agent approach. *Computers and Education: Artificial Intelligence, 3*, Article 100057. https://doi.org/10.1016/j.caeai.2022.100057

Willey, H., Howard, M., & Hutchinson, G. (2009). *SPARK-plus self and peer assessment guide v 1.7*. UTS-Sydney. https://sparkplus.com.au/

Wolterinck-Broekhuis, C. H. D., Poortman, C. L., Schildkamp, K., & VisscherKey, A. J. (2024). Key stakeholder voices: Investigating student perceptions of teachers' use of assessment for learning. *Educational Assessment, Evaluation and Accountability.* https://doi.org/10.1007/s11092-024-09428-7

Yan, Z., & Boud, D. (2022). Conceptualising assessment-as-learning. In Z. Yan & L. Yang (Eds.), *Assessment as learning. Maximising opportunities for student learning and achievement* (pp. 11–24). Routledge.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X. (2023). ChatGPT user experience: Implications for education. *SSRN*. https://doi.org/10.2139/ssrn.4312418

Zou, D., Xie, H., & Wang, F. L. (2023). Effects of technology enhanced peer, teacher and self-feedback on students' collaborative writing, critical thinking tendency and engagement in learning. *Journal of Computing in Higher Education, 35*(1), 166–185. https://doi.org/10.1007/s12528-022-09337-y

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

María Soledad Ibarra-Sáiz[1,2] · Miguel Ángel Gómez-Ruiz[1,2] · Antonio Balderas[3] · Gregorio Rodríguez-Gómez[1,2]

✉ Gregorio Rodríguez-Gómez
gregorio.rodriguez@uca.es

María Soledad Ibarra-Sáiz
marisol.ibarra@uca.es

Miguel Ángel Gómez-Ruiz
miguel.gomez@uca.es

Antonio Balderas
antonio.balderas@uca.es

[1] EVALfor Research Group, Facultad de Ciencias de la Educación, Universidad de Cádiz, Campus Puerto Real, 11519 Puerto Real, Cádiz, Spain

[2] UNESCO Chair in Evaluation & Assessment, Innovation and Excellence in Education, Universidad de Cádiz, Cádiz, Spain

[3] SPI&FM Research Group-Software Process Improvement and Formal Methods, Universidad de Cádiz, Puerto Real, Spain