Methodology for Quality Assurance of Educational Software

Norka Elvira Roca Ducasse¹, Walfredo González Hernández²

¹ University of Matanzas
² Universidad Central de Las Villas

Funding: Ministerio de Educación Superior, Gobierno de Cuba

Potential competing interests: No potential competing interests to declare.

Abstract

Evaluating the quality of educational software is a priority given the number of educational systems currently being produced. The article carries out a documentary analysis to search for documents that have addressed these elements. The results address the main documents that have been obtained. Subsequently, a methodology containing a system of metrics to evaluate the quality of educational software is proposed.

Keywords: educational software, methodologies, quality, quality metrics.

Introduction

The quality of software products is desired by both developer organisations and users. As expressed by Zhi et al. (2023) the pursuit of quality is the main guideline of organisations that influences all development processes and strengthens them with high standards. This ensures continuous improvement and evolution in the organisation. In a wide range of literature (López et al., 2022; Mishra & Otaibi, 2020) the quality of a product is associated with the fulfilment of a customer's requirements, which is why the customer's criteria are fundamental for its evaluation. Therefore, when software quality is addressed, it refers to a product that can meet the expectations of both teachers and users and that is free of errors.

To Bahamdain (2015) in open systems quality frameworks there are three processes that ensure the quality of the system: error detection, error checking and solution checking. While Naqvi et al. (2020) states that security, usability, accessibility, reliability, privacy, and sustainability are important characteristics of quality. There is widespread recognition that product quality depends on the quality of evidence (Barraood et al., 2022; Salimbeni et al., 2023). This article adds that audits are also essential as they verify the quality of the documentation resulting from the processes. The documentation has the fundamental function of describing the software processes and in them different problems can be detected during the development process.

For other authors, quality assurance is defined as a "... planned and systematic plan to assure management that defined
standards, practices, procedures and methods are applied in the processes” (Naik & Tripathy, 2008, p. 10).

In the case of educational software, its main objective is to achieve the educational purpose for which it was developed, which is the quality requirement it must fulfil. For this reason, quality assurance actions must be oriented towards the review and testing of the fundamental processes that enable the fulfilment of this objective. In the case of educational software, there is a wide range of typologies, including serious games, tutorials, reviewers, among others, in which work is done on their quality (Gauthier & Jenkinson, 2018; Pauliková et al., 2016), but this analysis is carried out separately. They are also considered as part of educational multimedia software and websites (Sánchez et al., 2016; Torres Kompen et al., 2019) which are not always programmed or are sometimes part of platforms such as virtual spaces. Therefore, in this article it is assumed that all of them constitute digital didactic resources that are developed to fulfil an educational objective.

Different methodologies are used during the development of digital learning resources (Medina Chicaiza & González Hernández, 2019) adapted to their conditions. This makes it necessary to unify all the processes and assume a general methodology that contains all the possible actions that can be executed. This article assumes the methodology proposed by Medina Chicaiza and González Hernández (2019) which is based on the generic model of software development (Karambir & Sharma, 2016; Pressman & Lowe, 2013), in the methodology proposed by Llerena Ocaña and González Hernández (2019) and in the extended methodology proposed by Cataldi et al. (2006). The assumed methodology contains 8 stages: Diagnosis of the current situation; Preparation of the initial conditions for implementation; Implementation; Quality assurance; Support assurance; Maintenance or recycling; Withdrawal; Evaluation of the previous actions of the methodology. The methodology proposes two fundamental moments for quality management, stages 4 and 8, which are transversal to the whole process according to the authors of the methodology (Medina Chicaiza & González Hernández, 2019). However, in this case these actions have two shortcomings: the first is that they are oriented towards virtual courses and the second is that they are not sufficiently detailed to be applied to any general teaching resource.

Methodology

The use of documentary analysis as a research method is becoming more and more common (Jiménez Vargas et al., 2017; Zaragoza Vega & Gutiérrez Pérez, 2019). Table #1 below summarises the most commonly used stages in the literature. The last stage that has been added is a proposal of this article as it should propose a solution if a research problem has been found as is the case.

| Table 1. Phases of desk research. |
## Stage One

A search was carried out in the Web of Science at www.sciencedirect.com with the terms (educational) and (software) and (quality) in the title of the article, obtaining 2 documents (https://www.sciencedirect.com/search?title=educational%20software%20quality). Another search was also structured with the following terms (education) and (software) and (quality), obtaining 3 documents (https://www.sciencedirect.com/search?title=education%20software%20quality). The same descriptors were used for the scielo search which yielded a single document as can be seen in the link https://search.scielo.org/?q=&lang=pt&count=15&from=0&output=site&sort=&page=1&q=%28ab%3A%28Educational%29%29+AND+%28ab%3A%28Software%29%29+AND+%28ab%3A%28Quality%29%29&lang=pt&page=1&q=%28ti%3A%28Educational%29%29+AND+%28ti%3A%28Software%29%29+AND+%28ti%3A%28Quality%29%29&lang=pt&page=1. The papers obtained from sciencedirect address the issue for the English language while others focus on some of the technological aspects such as reusability. Although there is a great deal of research into educational software (Tamim et al., 2021) generally focus on a single type of software such as e-learning courses (Medina et al., 2021), instructional design (Wang et al., 2021) or of didactic materials in online training.

## Stage Two

The search engines used yielded a total of 10 documents in which there is no general analysis of the processes necessary to guarantee the quality of educational software. This affirms the need to establish a methodology for the quality management of digital learning resources.

## Stage Three

Each of the documents found were inserted into a customised digital library that was processed with the EndNote 20 system. This system allows each of the documents and their metadata to be managed in a way that allows them to be read and cited in Word documents such as this article. The rest of the stages of the documentary analysis will be described in the remaining parts of the article.

## Result

Source: Authors' elaboration.
Phase One - Diagnosis of the educational software domain

The phase synthesises a system of actions to evaluate the diagnostic process carried out to obtain the current situation of the level of development that teachers have in order to develop their own digital didactic resources. It is important to assess the relevance of the teachers’ training in educational technology to enable them to develop their digital teaching resources. Emphasis will be placed on diagnostic instruments on the knowledge and use of free tools that allow them to obtain quality products without having to programme.

Objective: To evaluate the process of diagnosing the current state of the art of educational software development.

Here we talk about functional and non-functional requirements and decision making about the need for programming or not.

The actions to be carried out in this phase are detailed below:

- Assessment of the congruence between the theoretical assumptions about the educational processes to be computerised and the instruments for diagnosing the development of information competences in teachers and students that will enable them to appropriate the technological resources needed to develop digital teaching resources.
- Evaluation of the instruments to determine the characterisation of the teachers' knowledge of the contents of their teaching and the contents of the syllabus.
- Confirmation of the validity of the instruments obtained for the elaboration of instruments for the development of the teachers' performance in the development of virtual courses and the contents of their subject.
- Review of the process for determining the diagnosis of the software ecosystem of the institution: it is important to know the network architecture with its equipment and software that support the online course platforms, the speed of information transfer between the different channels, the levels of connectivity of the resources and how these are integrated to support the spaces must be diagnosed with load and stress tests.
- Evaluation of the diagnostic process of teachers' knowledge of blended learning.
- Evaluate the weaknesses and strengths detected by their correspondence with the reality obtained from the application of the instruments.

Second Two: Assessment of the management of the school context to ensure the development of the digital learning resources

Objective: To evaluate the management of the school context in order to solve the inadequacies detected in the previous phase to achieve the desired purpose.

This phase deals with the evaluation of the actions that have been taken to ensure the fundamental processes underpinning the development of digital learning resources. Each of the actions undertaken is oriented towards the review of documents and the intensive use of checklists to verify whether the actions were fulfilled or not. Each unfulfilled action...
will be assessed with the team to analyse its causes. This stage leads to the design of the fundamental processes to mitigate the weaknesses detected during the diagnosis and, although it is not possible to detect all the deficiencies detected beforehand, it is possible to establish some general actions to solve the problems detected in other investigations. The following actions are planned for the development of this stage:

- Evaluation of the use of some methodology or model that allows them to represent the components of the digital learning resource to be implemented in the next phase.
- Evaluate the application of the methods in order to identify shortcomings and potentialities.
- Evaluation of socialisation in workshops, seminars and meetings based on the number of interventions among the total number of participants and that all teachers have intervened at least once.
- Evaluation of the depth of the questions asked in the workshops, seminars and meetings about the contents to be dealt with in the education systems.
- Evaluation of the efficiency of the communication established between teachers when exchanging about the tools and technologies to be used.
- Evaluation of the effectiveness of actions to identify interdisciplinary relationships in subjects.
- Evaluation of the effectiveness of teachers’ cooperative work by knowledge areas, avoiding the creation of obstacles in communication.
- Evaluation of training actions for the actors involved in the development of the educational software, if necessary.
- Audit of the files obtained from previous iterations to discover shortcomings in the fulfilment of the corresponding actions and use them as examples (eliminating any personal data that could identify the persons involved) for the other actors involved in the process.
- Develop case studies with successful actors and actions to enable them to observe the ways of acting and ways of working that have led to success.
- Verification of the correct use of the artefacts proposed in the methodology or model used.
- Audit of the documentation generated from the introduction of the selected methodologies or models.

In this phase, each of the actions is documented through the preparation of the minutes of the meetings between teachers for their improvement, as well as the levels of satisfaction and dissatisfaction that they have. As a first prototype, the images constructed by the teachers on the structure that their virtual course should have are obtained as a first prototype.

Phase Three: Evaluation of the implementation processes of the digital learning resources

Objective: To evaluate the development process of digital teaching resources achieved by teachers.

The phase addresses the fundamental actions to achieve quality in digital learning resources with emphasis on the evaluation of collective actions and communication processes. The use of appropriate artefacts and the generation of the necessary documentation to achieve a climate of teamwork is important. This should help to ensure that the objectives of the development group are taken into account over and above the individual aspects that exist. For the development of this phase it is important to structure the following actions:
• Evaluate the effectiveness of the calls for teachers to join the educational software implementation process.

• To assess the structure, content and relevance of the knowledge to be included in the digital teaching resources proposed by Marciniak and Cáliz Rivera (2021).

• Assessing the effectiveness of the choice of knowledge units.

• Assessment of the suitability of the teachers in charge of each group to enable the correct orientation in the face of doubts and questions that may arise.

• Assessment of the suitability of the tools to be selected depending on the didactic objective to be achieved, in particular to eliminate programming that is complex for teachers.

• Evaluate the implementation of

• Evaluation of interoperability standards between digital learning resource development systems.

• Determination of the correspondence between the visual aspects of the digital learning resource, their need in relation to the pedagogical objective and the hardware requirements they need.

• Verify that the digital learning resources manage to give adequate feedback to learners during the interaction.

• Validate cultural elements that may be offensive or harmful to any of the students in any of the components of their own or others' digital learning resources.

• Evaluation of the correspondence between the digital learning resources and the proposed educational objective.

• Verify the usability, navigability, effectiveness, performance and reliability of each digital learning resource through stress, load, usability and navigability tests.

In case the proposal of the managers is to develop a system of digital didactic resources then it is necessary to evaluate the transit through the corresponding sub-phases in Llerena-Ocaña and González Hernández (2020) taking into account as an objective: To verify the efficiency and effectiveness of the actions proposed in the methodology for obtaining a system of digital didactic resources. To this end, the following actions are proposed:

• Unwrap the views of each content that will be shown to users;

• Verify compliance with the system approach in the development of digital learning resources, including the following aspects:

  • Use of similar visual resources in similar situations in learner interaction with digital learning resources.

  • Standardisation of virtual environments in the resources for similar tasks such as learning new content, exercising content already learned, online workshops to be held, video chats, among others.

  • Using the same sequences of screens in similar cases allows the learner to concentrate on learning the content and not on the sequence of screens to go through.

• Evaluate role-based access management as necessary during their transit from digital resource to digital resource.

• Execution of traceability tests of the results of previous digital learning resource systems to determine the previous difficulties of learners.

Phase Four: Strengthening support for digital learning resources
Objective: To establish actions to evaluate the assurance process necessary for the support of digital teaching resources developed by teachers.

Accessibility is one of the most desired qualities of digital learning resources because the didactic function for which they were designed can only be developed through interaction with learners. To achieve this, it is essential that there is effective feedback between the developers and the users of digital learning resources. In the case of the methodology developed by Medina Chicaiza and González Hernández (2019) these digital didactic resources are developed by the teachers so the feedback processes are not complex and are obtained from the direct interaction between the personal components of the process. The actions to be executed are:

- Verify the production of new digital teaching resources depending on the needs of the institution, students or teachers.
- Check whether the digital learning resources that have been downloaded are consistent with the rest of the resources produced in the institution.
- Validation of the ways of collecting criteria about the digital didactic resources that are implemented as part of the proposed educational media to be used during the teaching-learning process.

Phase Five: Evaluation of the continuous improvement of digital learning resources

Objective: To establish evaluation actions for the improvement processes of online courses that allow them to be adapted to changes in the environment.

Transformations characterise educational processes and digital learning resources are inherent to it. Therefore, digital learning resources are in constant transformation as a result of the tension between educational objectives, the developmental level of the learners and the content to be learned. Every transformation in digital learning resources that tries to balance the above mentioned triad must be documented and generate new resources with quality. To achieve this it is important:

- Evaluation of the establishment of strategies for modifying digital learning resources based on school transformations leading to a change in the initial requirements.
- Assessment of the efficiency and effectiveness of the actions implemented in order to obtain the necessary modifications.
- Audit of the documentation of the necessary transformations and their adequacy to the modifications that gave rise to them from the implementation of the actions proposed in the previous stages.
- Tests of the transformations carried out on the digital teaching resources, with emphasis on unity, regression, integration and the other transformations already proposed in previous actions.
- Audit of the collection of information from students on their satisfaction with the use of digital learning resources, particularly in terms of transparency in the application of the IAdov technique.

Phase Six: Evaluation of digital learning resources withdrawal actions
Objective: To establish the necessary actions for the disposal of digital learning resources when they are no longer needed.

Functional requirements respond to an educational need that can be expressed in the curriculum design of any educational level. Transformations in educational design lead to transformations in the configuration of didactic components during planning and implementation from the course to the classroom. Digital learning resources must comply with the educational designs for which they have been created and when it is not possible to perfect them they must be withdrawn.

Actions to be taken:

- Evaluation of the causes that led to the withdrawal of the digital learning resource(s).
- Assessment of the selection of the digital teaching resources to be eliminated and their relationship with the rest of the components of the teaching-learning process.
- Audit of the actions taken for the disposal of digital learning resources using software withdrawal templates.
- Conduct recall tests if the digital learning resource was embedded in a tightly coupled digital learning resource ecosystem.

Phase Seven: Assessment of the phases proposed above and their contribution to quality

Objective: To determine a system of tasks that will make it possible to obtain the relevant evaluations that will make it possible to correct the failures that may occur and that threaten the educational objectives of students and teachers.

Each of the proposed actions aims to detect the failures in each of the actions described above, which is why this phase is transversal in the methodology. To achieve this, the following is proposed:

- Establish the most effective channels of communication in full confidence that allow teachers to expose strengths, weaknesses and opportunities to reach consensus at each stage in order to detect failures.
- Collective critical reflection on the successes and failures that have occurred throughout the process that can enhance the team’s experience in these development processes.
- Implement an assurance system based on a quality assessment metric for virtual courses proposed in the literature (Medina et al., 2021).

However, any methodology must contain the ways in which the prototypes resulting from its actions will be evaluated. We agree with Tomas et al. (2013) and MEJIAS (2013) when they state that “Establishing metrics for measuring the quality of the software product is a basic piece of software quality control. A system for measuring the quality of a software product is efficient, has a high level of automation and allows for frequent use without excessive use of time” (p. 246). It is therefore important to consider a metric for measuring the quality of educational software.

For this purpose, we take as a basis the metric proposed by Medina et al. (2021) which measures the quality of virtual courses and can be adapted to digital teaching resources. For these authors, the metric is a sum function of attributes that
is multiplied by a weight factor; however, the qualities of a software are decomposed into factors, attributes and measurement criteria (Pressman, 2011). Taking this criterion into account, the following general metric is obtained:

\[
	ext{Quality} = \sum_{i=1}^{A} P_i \cdot \sum_{j=1}^{F} P_{ij} \cdot \sum_{k=1}^{CM} c_{mk} \]

where \( A \) is the number of attributes, \( F \) the number of Factors that each attribute possesses, \( CM \) is the number of measurement criteria that belong to each attribute, \( P_i \) and \( P_{ij} \) are the weights of the attributes and factors designated to them by each customer in accordance with the organisation, and \( c_{mk} \) are the measurement criteria associated with each attribute.

The main advantage of the metric obtained is that it does not depend on the attributes, factors or measurement criteria proposed by the developers. It can be used in any development process of any digital didactic resource, although it is recommended to unify criteria in terms of measurement criteria and weights that allow the metric to be coherent. In the same way, it can be used to evaluate the quality of any digital didactic resource as long as the factors, attributes and measurement criteria are clearly defined. This article proposes the use of the following scale for the measurement criteria, as it has proven its effectiveness in several research studies (Espín Andrade & González, 2000; Llerena-Ocaña & González Hernández, 2020): true (1), almost true (0.9), quite true (0.8), somewhat true (0.7), more true than false (0.6), as true as false (0.5), more false than true (0.4), somewhat false (0.3), quite false (0.2), almost false (0.1), false 0.

When determining the metrics without a system of factors, attributes and measurement criteria for educational software (an undertaking that would make this article very long) it is important to translate the values obtained into a scale that is understandable to the evaluator. In order to do so, the first thing the quality measurer has to fix are the maximum and minimum values of the range of the quality measurement scale \([VM_{\text{Mínimo}}, VM_{\text{Máximo}}]\). The maximum values of the proposed scale (1) are substituted to determine the \( VM_{\text{Máximo}} \) and the \( VM_{\text{Mínimo}} \) is equal to zero. When these are obtained, \( VM_{\text{Mínimo}} \) and \( VM_{\text{Máximo}} \) the equivalent scale can be calculated taking into account the values the assessor uses and the quality it represents. The first thing is that the evaluator must declare the number of intervals that his scale has and it will be named \( \text{NumInterval} \). To obtain the values of each interval, the ranges of each interval must be calculated given by the expression \( VM_{\text{Máximo}}/\text{NumInterval} \) and is denoted by \( \text{ValInterval} \). Having all these values we obtain that the first interval is \( (VM_{\text{Mínimo}}, VM_{\text{Máximo}}) \), being \( VM_{\text{Máximo}} = VM_{\text{Mínimo}} + \text{ValInterval} \) which in the case of this investigation would be \( (0, \text{ValInterval}) \). In the case of the second interval, it is necessary to determine the minimum of the interval, and for this we calculate \( VM_{\text{Mínimo}} = VM_{\text{Máximo}} + \frac{1}{10^n + 1} \), with \( n \) being the number of decimals of \( VM_{\text{Máximo}} \), in the case of the maximum we add again \( \text{ValInterval} \) to \( VM_{\text{Mínimo}} \). To calculate the minimum value of the last interval we would use:

\[
VM_{\text{Mínimo}} = VM_{\text{Máximo}} + \frac{1}{10^n + 1}
\]

and the value of the maximum would be \( VM_{\text{Máximo}} \).

The proposed evaluation system makes it possible to quantify such a subjective quality of educational software as quality and allows measurements to be established on scales recognised by implementers without having to learn new ones. This reduces the learning time of measurement processes while at the same time it does not set any of the factors, attributes and measurement criteria a priori, allowing each organisation to calculate with its own factors in mind. This provides measurement flexibility for development organisations.
Conclusions

The application of the documentary analysis made it possible to determine the shortcomings in the bibliography on the quality of educational software. The existence of documents, metrics and methodologies for the case of different educational software that evaluate specific elements for each of them is confirmed.

The proposed methodology contains the essential elements for the evaluation of the quality of an educational software. It was structured taking into account a system of phases and actions to fulfil its objective. Within the methodology, metrics were obtained that allow the evaluation of quality from a quantitative perspective.

Bibliography


