Cenas Educacionais

CONTINUOUS FLOW

e-ISSN: 2595-4881

MAPHYSE: MODEL FOR ASSESSING PHYSICS SIMULATORS FOR EDUCATIONAL PURPOSES

MAPHYSE: MODELO PARA LA EVALUACIÓN DE SIMULADORES DE FÍSICA CON FINES EDUCATIVOS

MAPHYSE: MODELO PARA AVALIAÇÃO DE SIMULADORES DE FÍSICA PARA FINS EDUCACIONAIS

Manuel Joaquim Silva de Oliveira [1]
Roges Horácio Grandi [2]
Raquel Salcedo Gomes [3]
Leandro Krug Wives [4]
José Valdeni de Lima [5]

Manuscript received on: April 13, 2023.
Approved on: December 13, 2023.
Published on: February 27, 2024.

Abstract
Simulators play a crucial role in the learning process of Physics, especially in institutions with limitations in experimental equipment, measurement instruments, or materials. The ability of students to use simulations to verify if the results match the expected facilitates their understanding of concepts. In this context, it is essential to assess whether simulators meet the necessary requirements for learning. This article presents MAPHYSE, a Model for Evaluating Physics Simulators for Educational Purposes, developed based on pedagogical requirements of ISO/IEC 25010. The use of this standard aimed to establish clear pedagogical objectives, identifying essential characteristics such as functional adequacy and usability. The

[1] Doctorate in Information Technology in Education from the Federal University of Rio Grande do Sul. Professor at the Pedagogical University of Maputo. Member of the Research group Active Methodologies Supported by Digital Technologies.
ORCID: https://orcid.org/0000-0003-1918-5351 Contact: moliveira@up.ac.mz / jocasiloliveira79@gmail.com
[2] Doctorate student in Educational Informatics and Master in Computer Science from the Federal University of Rio Grande do Sul. Systems Architect at the Federal Data Processing Service. Member of the research group Learning Trajectories in Ubiquitous Hyperdocuments.
ORCID: https://orcid.org/0000-0001-9987-379X Contact: roges.grandi@gmail.com
[3] Doctorate in Informatics in Education from the Federal University of Rio Grande do Sul. Professor in the Postgraduate Program in Informatics in Education at the Federal University of Rio Grande do Sul. Member of the research group Learning Trajectories in Ubiquitous Hyperdocuments.
ORCID: https://orcid.org/0000-0001-9497-513X Contact: raquel.salcedo@ufrgs.br
[4] Doctorate in Informatics in Education from the Federal University of Rio Grande do Sul. Professor in the Postgraduate Program in Informatics in Education at the Federal University of Rio Grande do Sul. Member of the research group Learning Trajectories in Ubiquitous Hyperdocuments.
ORCID: https://orcid.org/0000-0002-8391-446X Contact: leandro.wives@ufrgs.br
[5] Doctorate in Informatics from Université Joseph Fourier. Professor in the Postgraduate Program in Informatics in Education at the Federal University of Rio Grande do Sul. Member of the research group Learning Trajectories in Ubiquitous Hyperdocuments.
ORCID: https://orcid.org/0000-0002-7266-4856 Contact: valdenilima@gmail.com

model includes evaluation criteria and measurable metrics, iteratively adjusted based on tests and feedback. Comprehensive documentation and user guides were fundamental, along with training processes for effective implementation. The implementation of the model in educational contexts, with continuous monitoring, aims to create a robust tool for comprehensive assessment of the quality of educational simulators. The validation of the model involved the evaluation of two simulators, indicating that MAPHYSE offers a wide range of criteria to assess the quality of simulators, being a useful tool for schools, teachers, and developers in decision-making and continuous improvement of educational tools.

Keywords: Physics simulators; Quality reference models; ISO/IEC 25010 Standard; Software quality assessment.

Resumen

Los simuladores juegan un papel crucial en el proceso de aprendizaje de la Física, especialmente en instituciones con limitaciones en equipos experimentales, instrumentos de medición o materiales. La capacidad de los estudiantes de utilizar simulaciones para verificar si los resultados coinciden con lo esperado facilita su comprensión de los conceptos. En este contexto, resulta fundamental evaluar si los simuladores cumplen los requisitos necesarios para el aprendizaje. Este artículo presenta MAPHYSE, un modelo para la evaluación de simuladores de física con fines educativos, desarrollado con base en los requisitos pedagógicos de la norma ISO/IEC 25010. El uso de esta norma tuvo como objetivo establecer objetivos pedagógicos claros, identificando características esenciales como la adecuación funcional y la usabilidad. El modelo incluye criterios de evaluación y métricas medibles, ajustadas iterativamente en función de pruebas y comentarios. La documentación completa y las guías de usuario fueron fundamentales, junto con los procesos de capacitación para una implementación efectiva. La implementación del modelo en contextos educativos, con seguimiento continuo, tiene como objetivo crear una herramienta robusta para la evaluación integral de la calidad de los simuladores educativos. La validación del modelo implicó la evaluación de dos simuladores, indicando que MAPHYSE ofrece una amplia gama de criterios para evaluar la calidad de los simuladores, siendo una herramienta útil para escuelas, docentes y desarrolladores en la toma de decisiones y mejora continua de herramientas educativas.

Palabras clave: Simuladores de física; Modelos de referencia de calidad; Norma ISO/IEC 25010; Evaluación de la calidad del software.

Resumo

Os simuladores desempenham um papel crucial no processo de aprendizagem da Física, especialmente em instituições com limitações em equipamentos experimentais, instrumentos de medição ou materiais. A capacidade dos alunos de utilizar simulações para verificar se os resultados correspondem ao esperado facilita sua compreensão dos conceitos. Nesse contexto, é fundamental avaliar se os simuladores atendem aos requisitos necessários para a aprendizagem. Este artigo apresenta o MAPHYSE, um Modelo de Avaliação de Simuladores de Física para Fins Educacionais, desenvolvido com base em requisitos pedagógicos da Norma ISO/IEC 25010. O uso dessa norma buscou estabelecer objetivos pedagógicos claros, identificando características essenciais como adequação funcional e usabilidade. O modelo inclui critérios de avaliação e métricas mensuráveis, ajustados iterativamente com base em testes e feedbacks. Documentação abrangente e guias do usuário foram fundamentais, juntamente com processos de capacitação para implementação eficaz. A implementação do modelo em contextos educativos, com monitoramento contínuo, visa criar uma ferramenta robusta para avaliação abrangente da qualidade dos simuladores educacionais. A validação do modelo envolveu a avaliação de dois simuladores, indicando que o MAPHYSE oferece uma ampla gama de critérios para avaliar a qualidade dos simuladores, sendo uma ferramenta útil para escolas, professores e desenvolvedores na tomada de decisões e na melhoria contínua das ferramentas educacionais.

Palavras-chave: Simuladores de física; Modelos de referência de qualidade; Norma ISO/IEC 25010; Avaliação de qualidade de software.

Introduction

Understanding physical phenomena requires a high level of abstraction. Simulators (computational representations of real systems) can support the learning of such phenomena (Bastos; Wilkinson, 2010). They are usually based on deterministic semantic models, i.e., for the same inputs, it is expected the same outputs. As learning resources, they can arouse interest in complex phenomena and raise the level of abstraction. They help students pay attention to the physical principles involved in the phenomenon and not just to mathematical procedures. Developing activities with simulators involves several abstraction levels in a pedagogical strategy of four phases: motivation, research, formalization, and transference (Fernandéz, 2005, 2000).

However, there is a gap on models for evaluating simulators from the learning perspective. Evaluating pedagogical software is a complex process that demands an adequate strategy (Silva *et al.*, 2016). An approach that enables all-embracing software evaluations is to use a reference model such as ISO/IEC[6] 25010 (2011), which offers a broad set of characteristics and sub-characteristics organized around three quality perspectives: internal, external, and in use. Experts (in our case, teachers) evaluate simulators using external processes and users (students) the quality in use. Developers may receive feedback from them and start improvement cycles as part of internal quality processes.

[6] International Organization for Standardization/International Electrotechnical Commission.

In a systematic literature review[7], we did not find specific studies for assessing Physics simulators from the point of view of learning, although models for evaluating pedagogical software for other purposes were found. This fact motivated us to create the **Model for Assessing Physics Simulators for Educational Purposes (MAPHYSE)** based on external and in-use perspectives of the ISO/IEC 25010 Standard. The internal quality processes were not included because developers' involvement could inadvertently trend to positive assessments.

A process for evaluating MAPHYSE suitability was performed assessing three Physics simulators having it as the quality reference model. Based on the assessment processes' consistency and the produced reports, we concluded that the Model is adequate to assess Physics simulators for educational purposes. The mentioned reports can lead Physics teachers to a better understanding whether simulators are adequate for their teaching activities.

Theoretical Foundation and Related Works

It is of great value that students of Natural Sciences understand correctly and early phenomena, concepts, theories, laws, nomenclatures, units of measure, and representations following international norms, avoiding the abstraction of alternative concepts, also known as spontaneous. Learning Physics is an area full of difficulties, and it is common for alternative conceptions to limit the deep understanding of various domains within this discipline (Conde, 2021).

According to Leão and Kalhil (2015), alternative concepts are knowledge that, although not under scientific rigor, can evolve to more accurate concepts according to the theories and the laws that describe them.

If students formed alternative concepts into their cognitive structure, teachers should act to help them be able to perceive consistent meanings about phenomenology and science. Knowledge derived from conceptual evolutions

[7] This review was realized in December 2020.

expands the capacity to perform experiments and build appropriate conceptions about the natural, social, and technological environment. Teachers must pay attention to the complexities inherent to knowledge about nature, technology, and society and remember that students gradually build them as they develop their cognitive facets (SCHROEDER, 2007). Queiroz and Lima (2007) assert that new knowledge in science is produced by creative acts of imagination combined with rigorous but varied scientific research methods; its acquisition is problematic, never easy. According to Leão and Kalhil (2015), teachers mediate apprenticeship on concepts, discussing scientific knowledge, making deductions from formulas, understanding energy, talking about natural phenomena, and entering the world of sciences.

Low-quality simulators tend to have negative influences on learning. Assessment models based on ISO/IEC 25010 (2011) may help educational institutions, teachers, and students properly evaluate them from the perspective of learning objectives.

ISO/IEC 25010:2011 Standard

ISO/IEC 25010 (2011) is an international standard that defines a comprehensive set of characteristics and sub-characteristics that assess software quality from different perspectives. Using internal evaluation processes, developers may assess their products. External quality assessments are carried out by a group of specialists in the area and may include representatives from the development team. The quality in use is mainly assessed by users (students) and complemented by specialists (teachers).

The Standard arranges both internal and external assessment characteristics into 8 classes (Figure 1): Functional Suitability, Reliability, Usability, Performance Efficiency, Maintainability, Portability, Compatibility, and Security, each with its own set of sub-characteristics. The quality in use perspective has 5 groups: effectiveness, efficiency, satisfaction, freedom from risk, and context coverage, as shown in the hierarchical diagram (Figure 2).

Figure 1: Internal and external quality model. Source: ISO/IEC (2011).

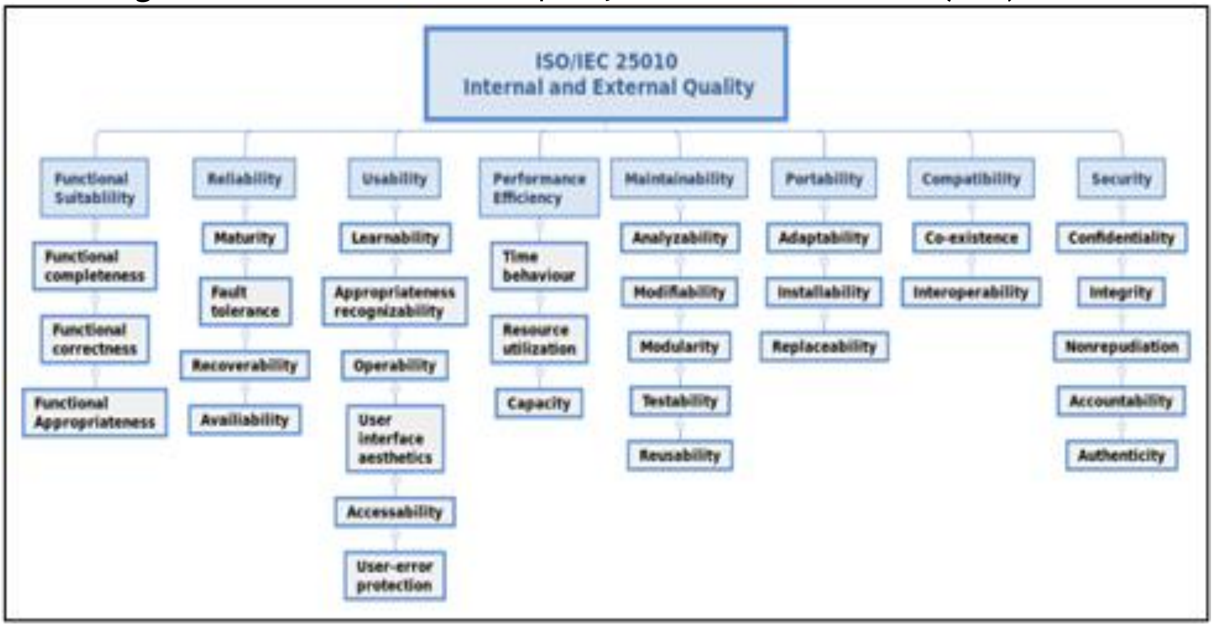
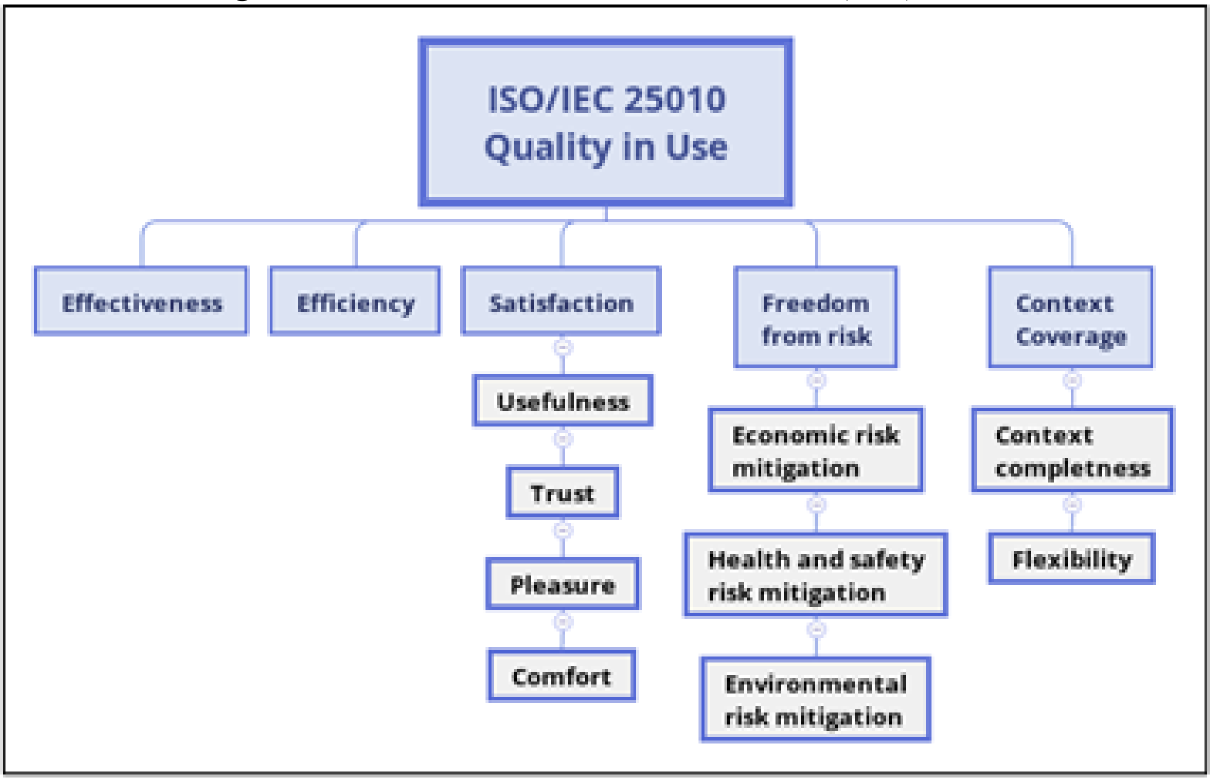


Figure 2: Quality in use model. Source: ISO/IEC (2011).



- Models for Physics Simulator Assessment

According to a particular set of criteria, a reference model for software assessment is a standard for measuring a software product's quality. A systematic literature review was carried out to answer the following research question: *How to estimate the quality of Physics simulators in an educational context using a mature and internationally recognized assessment model?* The context of this research question was defined by PICOC[8] criteria established (Petticrew and Roberts, 2008) as follows:

1. Population: Physics simulators.

2. Intervention: mixed-method assessment based on ISO/IEC 25010.

3. Comparison: quality characteristics.

4. Outcomes: mixed-method assessments of Physics simulators.

5. Context: teaching and learning processes of Physics.

The keywords were “25010”, “Physics” and “Simulator” in English, Portuguese, and Spanish. For *Google Scholar,* the keyword “Quality”, in the same languages, has been added to refine the search (Figure 3).

Figure 3: Search strings. Source: The authors.

```
Traditional academic indexers:

("25010" AND ("Physics" OR "Física") AND ("Simulator" OR "Simulador"))

Google Scholar:

("25010" AND ("Physics" OR "Física") AND ("Simulator" OR "Simulador") AND
("Quality" OR "Qualidade"))
```

The only inclusion criterion used was a study related to the quality assessment of physics simulators by ISO/IEC 25010. The exclusion criteria were: i) duplicate studies; ii) studies in languages other than those defined, and iii) studies not available for access. The publication date was not used as an exclusion criterion. In the preliminary selection, 31 studies were found: 16 in Google Scholar, 7 in Scopus, 7 in

[8] Population, Intervention, Comparison, Outcome, Context.

SpringerLink, and 1 in the ACM Digital Library. No studies have been found in El Compendex, the IEEE Digital Library, the ISI Web of Science, Science@Direct, and CAPES Journals.

Six studies were eliminated from Google Scholar by the criterion of duplicity. Scopus and SpringerLink found precisely the same studies, and these duplications were also eliminated. All papers found by Scopus and ACM were eliminated once the number "25010" did not refer to the ISO/IEC Standard.

All studies read from Google Scholar referred to ISO/IEC 25010. However, six of them do not deal with simulators or Physics. One paper only cited the Standard and simulators without developing the subject. Another study presented theoretical research on models, techniques, and instruments for assessing educational software. Although not directly dealing with the Physics area, the remaining two papers discussed ISO/IEC 25010 and simulators. These articles are presented below as related works.

- Related Work

Silva (2015) presented a Nursing Process Application System, which implements a realistic case simulator. ISO/IEC 25010 was used to guide the validation process carried out by a team composed of 10 nurses and 11 computer scientists.

Mathis *et al.* (2021) assessed the suitability of using Virtual Reality (VR) to evaluate the usability and security of real-world authentication systems. To do so, they conducted a replication study and created a virtual replica of CueAuth, a recently introduced authentication scheme, reporting results from: (1) an in-VR lab-based usability study evaluating user performance; (2) an online security study assessing the system's resistance to observation through virtual avatars; and (3) a comparison between their results and those previously reported in the real-world evaluation. The study's analysis indicated that VR can serve as a suitable platform for user-centered evaluations of real-world authentication schemes, but the utilized VR technology can modify the experiments.

Oliveira (2015) proposed a simulator for robotic systems as an alternative to reduce costs during development stages and to increase both the productivity and reliability. To assess the simulator quality, he asked users to evaluate of its attributes based on the 25010 Standard also using a 5-level Likert scale.

Siqueira (2022) acknowledges the fundamental role of validating simulators, and his research aimed to develop an evaluation model for the Virtual Tactical Simulator in the missile and rocket system operator course for officers. The proposal is based on fire support simulators used in the Brazilian Army. The study thoroughly examines the evaluation process of practical tests for Cadets of the Black Needles, highlighting the use of simulation as an integral part of this process.

Herpich *et al.* (2019) presented the MAREA evaluation model to assess mobile augmented reality educational approaches accordingly to: a) usability: based on ISO/IEC 25010 five dimensions; b) engagement: related to external and in-use qualities; c) motivation: related to the quality in-use and d) active learning: embracing effectiveness, challenge, feedback, security, and complexity dimensions. Benítez Guadarrama (2021) defines usability as a process of interaction between the subject and the object to carry out specific tasks or activities efficiently and effectively.

Soad (2017) proposed the Mobile Learning Evaluation method to evaluate mobile educational applications. Its quality model is divided into Technical, Pedagogical, and Social categories. The Technical Category is based on ISO/IEC 25010 internal/external quality characteristics. The Pedagogical and Social categories allow specialized assessment of functional adequacy and quality in-use regarding pedagogical and sociological requirements, respectively. A limitation of it is the focus on mobile applications. Since the functional adequacy is assessed in three categories (technical, pedagogical, and social), the teams responsible for the assessments must be careful of duplicity and possible conflicting results in different categories.

The mentioned models focus on different evaluation areas and have specific purposes regarding usability, security, simulation, mobile learning. MAPHYSE stands out for evaluating simulators of physical phenomena in educational processes based on a comprehensive approach to evaluation, including both teachers (experts) and students (users).

Development Methodology for Maphyse

The methodology for developing an assessment model for simulators, aligned with the principles of the ISO/IEC 25010 standard followed a structured approach. Initially, the model's objectives were defined to ensure the quality, usability, and effectiveness of educational simulators of physical phenomena. Subsequently, the requirements of ISO/IEC 25010 were studied and adapted to this specific context.

It was necessary to identify the critical quality characteristics for the effectiveness of these educational tools. Functional suitability and usability are examples of such characteristics. From there, evaluation criteria were defined, each accompanied by measurable metrics for an objective assessment.

The model was then constructed, organizing characteristics, criteria, and metrics in a structured manner. A scoring scale was established for each metric, providing a solid foundation for assessment. The model underwent tests using existing simulators, and feedback from software and education experts was collected.

Adjustments and refinements were made iteratively based on test results and received feedback. The model's documentation was comprehensive, including clear instructions on its application, along with user guides highlighting best practices and insights for improving simulators.

To ensure effective model’s usage, training was provided to those responsible for applying the model, ensuring a proper understanding of the evaluation criteria.

Model for Assessing Physics Simulators for Educational Purposes (Maphyse)

MAPHYSE is detailed in the following subsections according to the software characteristics defined by the ISO 25010 Standard. Each of its assessment items is associated with a sub-characteristic, which is identified by an unique reference and can be ranked in four levels. **Level 0 (not assessed)** means that the item has been disregarded in a specific assessment. **Level 1 (unsatisfactory)** is the lowest rating in an assessment. It means that the simulator did not fully meet the requirements or needs related to the item. **Level 2 (satisfactory with restrictions)** is the second ranking level. It means that the simulator has met the requirements with restrictions related to the item. **Level 3 (satisfactory without restrictions)** is the highest rating. It means that the simulator has met, without restrictions, the related requirements or needs.

Each assessment item has also a weight ranging from 1 to 3. It serves to emphasize, or not, the importance of the item in relation to the learning objectives. **Weight 1 (not very relevant)** means that a possible low assessment should not significantly compromise learning objectives. **Weight 2 (average relevance)** means that a possible low evaluation of the item may compromise, with median relevance, the learning objectives. **Weight 3 (very relevant)** means that a possible low assessment of the item will significantly compromise the learning objectives.

Multiplying the **Level by the Weight** generates the **item assessment rate**. A simulator that has one or more very relevant items (Weight 3) classified as unsatisfactory (Level 1) will have in its final evaluation a precautionary recommendation addressed to educational institutions and teachers, informing potential negative impacts on educational objectives. On the other hand, if all very relevant items are minimal or entirely satisfactory (Levels 2 or 3), the final evaluation will have a recommendation for use, stating that the simulator meets minimum requirements for its adoption in pedagogical activities.

- Internal Quality Assessment

Since the development team evaluates a simulator's internal quality and one of the objectives of MAPHYSE is to allow the recommendation or not of physics simulators in pedagogical processes, the use of internal vision is avoided as a precaution to possible conflicts of interest.

- External Quality Assessment

Functional suitability sub-characteristics – **functional completeness, correctness and appropriateness** (**S_{e1}..S_{e3}**) receive weight 3 for its relevance to the learning objectives (Table 1).

Table 1: Functional Suitability sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{e1}	$P_{e1} = 3$	Ability to simulate all the experiments that physical equipments can provide.
S_{e2}	$P_{e2} = 3$	Accurate calculations and graphical representations must be accurate.
S_{e3}	$P_{e3} = 3$	Simulations should empower students to achieve the learning goals.

Reliability items are listed in Table 2. **Maturity** (**S_{e4}**) receives weight 2, since it is relevant to continuity, especially during periods when students are accompanied by their teachers. **Fault tolerance, recoverability and availability** (**S_{e5}..S_{e7}**), although relevant, tend to compromise less than a low maturity (weight 1).

Table 2: Reliability sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{e4}	$P_{e4} = 0\|2$	Be **tFE** the failures presented during simulations. If **tFE <= 1%** then it is satisfactory without restrictions. If **1 < tFE <= 2%** then it is satisfactory with restrictions. If **tFE > 2%** then it is unsatisfactory.
S_{e5}	$P_{e4} = 0\|1$	Maintain a enough performance level when anomalous situations occur.
S_{e6}	$P_{e4} = 0\|1$	Regain operational normality when extreme destabilization situations occur.
S_{e7}	$P_{e4} = 0\|1$	Be operational and available when learning activities are performed.

Learnability (**S_{e8}**) is about the ability to understand the key concepts of the simulator to develop learning activities. Since it can influence the learning objectives, it is weighted 2. **Appropriateness Recognizability** (**S_{e9}**) measures students' recognition that the simulator effectively helps them appropriate the desired learning objectives. If they understand that it is not helping their learning, they will not feel encouraged to use it. For this reason, it also weights 2. **Operability – General** (**S_{e10}**) is the ease to perform tasks. **Operability – Language** (**S_{e11}**) ascertains the native language usage. **S_{e10}** and **S_{e11}** influence the learnability and the motivation for its effective usage (weight 2).

The **User Interface Aesthetics** (**S_{e12}**) assesses whether the simulator is considered pleasurable during the interactions (weight 1). **User-error Protection** (**S_{e18}**) measures the ability to prevent users from making basic operational errors that hinder processes. If the interface fails in these checks, learning activities may delay or be in any way difficult, but it will not prevent them from running (weight 1).

Accessibility deals with digital inclusion in pedagogical praxis – **visual, hearing or motor difficulties, autism, dyslexia (S_{e13}, S_{e17})**, among others. It is a constant and important challenge (NUNES, 2020; WHO 2011). Autistic people are favored by text and images rendered in a color spectrum that avoids very dark tones, very light or bright colors. For people with dyslexia, texts with left alignment, no underlined words, and no capital words favor reading. Texts, buttons, and other elements of screens favor people with the most different disabilities when written with adequate or adjustable size, direct, precise orientations, with denotative meaning, with fonts without serif, for making reading more understandable to most users (UFRGS, 2020; BRITTO et al., 2016; CAMPÊLO, 2013). Accessibility items were considered medium relevance or very relevant (Weights 2 or 3), given their importance for educational inclusion. As a standard, we adopted Weight 2, given the current state-of-the-art difficulty of Software Engineering, of producing effectively accessible physics simulators (Table 3).

Table 3: Usability sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{e8}	P_{e8} = 0|2	Allow key concepts to be easily understood enabling students to acquire skills to perform the experiments necessary for their learning.
S_{e9}	P_{e9} = 0|2	Help students to appropriate the desired learning objectives.
S_{e10}	P_{e10} = 0|2	Easiness to use in performing learning activities.
S_{e11}	P_{e11} = 0|2	Operation in the mother tongue.
S_{e12}	P_{e12} = 0|1	Pleasure and satisfactory interactions during learning activities.
S_{e13}	P_{e13} = 0|2|3	Needs of students/teachers with visual difficulties – blindness, low visual acuity, and color blindness in their respective mother tongues, and the mother tongues.
S_{e14}	P_{e14} = 0|2|3	Needs of students/teachers with hearing difficulties – deafness and low hearing capacity – in their mother tongues.
S_{e15}	P_{e15} = 0|2|3	Needs of students/teachers with motor difficulties.
S_{e16}	P_{e16} = 0|2|3	Needs of autistic students/teachers.
S_{e17}	P_{e17} = 0|2|3	Needs of dyslexic students/teachers.
S_{e18}	P_{e18} = 0|1	Prevention from doing basic operational errors that may hinder learning processes.

The **Performance Efficiency** is related to the simulator's ability to complete experiments with minimized resources used. The **Time Behavior (S_{e19})** classifies the degree to which the response times are satisfactory. If the response times are excessive, students may be discouraged about their use. For these reasons, it has weight 2. In MAPHYSE, the resource utilization is divided into **Network and Storage Resource Utilization (S_{e20}, S_{e21})** which measure, respectively, the degree of satisfaction with network bandwidth consumption and the storage space required by the simulator. Despite their respective relevance in efficiency, a higher network consumption and ample storage space usually do not compromise the experiments. Therefore, they receive weight 1. The **Capacity (S_{e22})** measures how much the simulator limits are enough to accomplish the learning objectives. For this reason, its weight 2 (Table 4).

Table 4: Performance Efficiency Sub-characteristics. Source: The authors

Ref.	Weight	Quality Requirements
S_{e19}	P_{e19} = 0|2	Be **t** the response time. If **t < = pTimeA** it is satisfactory; if **pTimeA < t <= pTimeA** it is satisfactory with restrictions; if **t > pTimeB** it is unsatisfactory.
S_{e20}	P_{e20} = 0|1	Be **b** the bandwidth consumed for transmitting an image. If **b <= pBandwidthA** it is satisfactory; if **pBandwidthA < b <= pBandwidthB** it is satisfactory with restrictions. If **b > pBandwidthB** then it is unsatisfactory.
S_{e21}	P_{e21} = 0|1	Be **s** the space to install. If **s <= pSpaceA** it is satisfactory; if **pSpaceA < s <= pSpaceB** it is satisfactory with restrictions. If **s > SpaceB** it is unsatisfactory.
S_{e22}	P_{e22} = 0|2	Computational limits should be enough to achieve learning objectives.

To assess **Maintainability**, it is usually required to access both code and technical documentation. Defect fixes are essential. If the simulator has a defect and takes time to fix or the fixes are unsatisfactory, learning objectives may be compromised. A cohesive and decoupled software architecture favors the analysis, development, and maintenance processes. Therefore, **analyzability, modifiability, modularity and testability (S_{e23} ..S_{e26})** receive weight 2. Possible contributions to other projects provided by **Reusability (S_{e27})** are important as a paradigm, but its absence does not directly compromise the simulator's educational objectives, receiving weight 1 (Table 5).

e-ISSN: 2595-4881

Table 5: Maintainability Sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{e23}	P_{e23} = 0|2	Acceptable effort/time to detect and debug errors when they occur.
S_{e24}	P_{e24} = 0|2	Acceptable effort/time to correct and evolve the simulator.
S_{e25}	P_{e25} = 0|2	Acceptable cohesive and uncoupling characteristics provided by the software architecture.
S_{e26}	P_{e26} = 0|2	Acceptable effort/time to perform regression tests.
S_{e27}	P_{e27} = 0|1	Acceptable ability to reuse parts of the code.

Adaptability (S_{e28}) allows students to use other alternatives for the simulations. However, its absence does not compromise educational objectives (weight 1). **Portability** addresses the ability to be transferred from one environment to another. The **Installability (S_{e29})** is important both for teachers and students to use the simulator on the devices they have. If they cannot use the simulator due to lack of installation capacity, it may compromise the desired educational objectives (weight 3). **Replaceability (S_{e30})** is crucial when you want to swap the execution platform. It becomes important when the new simulator must have the ability to perform the same pedagogical activities as the previous one, receiving weight 3 since it is directly related to educational objectives (Table 6).

Table 6: Portability sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{e28}	P_{e28} = 0|1	Adapt to other environments without additional settings.
S_{e29}	P_{e29} = 0|3	Be installable on the devices that teachers and students need.
S_{e30}	P_{e30} = 0|3	Be able to replace other simulators on the same devices used by teachers and students, providing the same learning opportunities.

The **Compatibility** is related to the simulator's ability to exchange information – **Interoperability-Import (S_{e32})** and **Interoperability-Export (S_{e33})** – and to perform its required functions while sharing the same environment. Import and export operations are important to enable the reuse of experiments in other simulators. Relevant, but if it does not have this capacity, it does not compromise primary educational objectives (weight 1). On the other hand, if a simulator destabilizes itself in the presence of other systems or, in reverse, destabilizes others – **Co-existence** sub-characteristic **(S_{e31})** – learning objectives can be compromised, receiving weight 2 (Table 7).

Table 7: Compatibility Sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{e31}	$P_{e31} = 2$	Perform their functions without negatively affecting the performance of other systems.
S_{e32}	$P_{e32} = 1$	Import information from other simulators in accordance with the established import requirements.
S_{e33}	$P_{e33} = 1$	Export information to other simulators in accordance with established export requirements.

Security handles protection characteristics for functions and data. When demanded, they are highly relevant to ensuring confidence in educational processes. For this reason, all safety items are weighed 3. **Confidentiality-Teachers (S_{e34})** is about avoiding unauthorized access to teachers' exclusive functions and data. **Confidentiality-Students (S_{e35})** also is about to prevent unauthorized access in the context of students. **Integrity (S_{e36})** addresses improper data modification during communications. **Nonrepudiation (S_{e37})** is about proving the authorship of actions performed. The **accountability (S_{e38})** of the ability to audit actions and **authenticity (S_{e39})** is the ability to provide mechanisms for identifying authors in institutional processes (Table 8).

TABLE 8: Security sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{e34}	$P_{e34} = 0\|3$	Avoid unauthorized access to functions and data of exclusive access to teachers.
S_{e35}	$P_{e35} = 0\|3$	Avoid unauthorized access to functions and data of exclusive access to students.
S_{e36}	$P_{e36} = 0\|3$	Avoid undue data changes during communications.
S_{e37}	$P_{e37} = 0\|3$	Be able to prove the authorship of actions.
S_{e38}	$P_{e38} = 0\|3$	Be able to trace actions.
S_{e39}	$P_{e39} = 0\|3$	Be able to authenticate users under the established authenticity requirements.

- Quality Assessment in Use

In use quality assessments can be useful to generate feedback that empowers developers to improve the simulator in subsequent evolutionary cycles. These improvements may include bug fixes, refinements, changes, and characteristic inclusions. Alone, the quality in-use is not enough to recommend the use or non-pedagogical use of the simulator. It is necessary to carry out, jointly, an evaluation of internal quality, external quality, or both to assess functional adequacy.

Effectiveness (**S_{u1}**) measures how much the simulator effectively helps learners to achieve learning objectives, from their own points of view. Since the achievement of significant learning objectives is achieved, it has weight 3. **Response Time Efficiency** (**S_{u2}**) measures satisfaction about the response time (Table 9).

Table 9: Effectiveness and Efficiency Sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{u1}	**P_{u1} = 0|3**	Effectively assist apprentices in achieving the established learning objectives.
S_{u2}	**P_{u2} = 0|1**	Learners should be satisfied with the response time of the simulator characteristics.

Satisfaction is a characteristic related to needs. Thus, the **Usefulness** (**S_{u3}**) should be measured in the context of learning satisfaction when they achieve the planned learning objectives. **Trust** (**S_{u4}**) supports confidence that the simulator helps students achieve their planned learning goals. **Pleasure** (**S_{u5}**) measures how much students find the simulator pleasurable to use, and how satisfying the interactions are. The **Comfort** (**S_{u6}**) item is a measure of physical comfort experienced by students when using the simulator. Satisfaction is important to feel stimulated to continue using the simulator. However, since it is a measure of contentment and not of achieving educational objectives themselves, its evaluation items have weight 2 (Table 10).

TABLE 10: Satisfaction sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{u3}	**P_{u3} = 0|2**	Students must be satisfied when they achieve their planned learning objectives with the help of the simulator.
S_{u4}	**P_{u4} = 0|2**	Students must be trusted that the simulator helps them achieve their planned learning objectives.
S_{u5}	**P_{u5} = 0|2**	Students must consider the interactions with the simulator pleasurable.
S_{u6}	**P_{u6} = 0|2**	It is necessary that students consider that interactions with the simulator provide physical comfort.

The **Freedom from Risk** is related to the minimization of potential risks. **Economic Risk Mitigation** (**S_{u7}**) is multifactorial. Risks are mitigated, for example, when: i) the use license is out of costs; (ii) training costs are acceptable or also absent; iii) when the simulator can be used for a long time to compensate for investments made. This item can weigh 2 or 3, according to the capacities and needs of each educational institution. By default, we set it to 2. **Health and Safety Risk Mitigation (S_{u8})** aims to minimize health risks by developing ergonomic interfaces. For instance, instead of offering only options for moving graphics by mouse, keyboard options could also be offered on desktop computers. **Environmental Risk Mitigation (S_{u9})** is related to sustainability influenced by energy efficiency and resource use (KOÇAK *et al.*, 2015). Risk mitigation is relevant in any project, including the implementation of educational programs. It is not, however, as expressive as the effective achievement of objectives. Thus, the sub-characteristics of absence of risks (Table 11) receive weight 2.

TABLE 11: Freedom from Risk sub-characteristics. Source: The authors.

TABLE 10: Satisfaction sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{u7}	P_{u7} = 0|2	If the license to use is free: satisfactory; if paid once, with reasonable price: satisfactory with restrictions; if paid monthly or once with an unacceptable price: unsatisfactory.
S_{u8}	P_{u8} = 0|2	Ergonomics should minimize the risk of causing repetitive strain injuries.
S_{u9}	P_{u9} = 0|2	The consumption of network resources and electricity resources should be minimized.

Context Coverage deals with ability to use the simulator effectively and efficiently in the context of specified use. **Context Completeness (S_{u10})** measures coverage to minimum learning contexts for achieving educational objectives. **Flexibility (S_{u11})** complements **S_{u10}** by measuring coverage in extra-contexts. The simulator must empower its use effectively in the minimum contexts desired to achieve educational objectives to offer different learning opportunities. For this reason, it receives weight 3. Flexibility, although relevant for providing alternatives to students, which can stimulate creativity and lateral thinking, is not as essential as achieving the planned educational objectives, thus receiving weight 2 (Table 12).

TABLE 12. Context Coverage sub-characteristics. Source: The authors.

Ref.	Weight	Quality Requirements
S_{u10}	P_{u10} = 0|3	Offer contexts of experimentation necessary for the achievement of basic learning objectives.
S_{u11}	P_{u11} = 0|2	Offer other contexts of experimentation besides those necessary for the achievement of basic learning objectives.

- External Quality Assessment Rate

To assess the quality from the external view, the assessment team assembles an equation with the criteria that it considers relevant or can evaluate (Figure 4).

Figure 4: Calculation of the External Quality Rate (TQe). Source: The authors.

$$SQ_e = \sum_{i=1}^{39} P_{ei}S_{ei} = P_{e1}S_{e1} + P_{e2}S_{e2} + P_{e3}S_{e3} + \cdots + P_{e39}S_{e39}$$

$$PQ_e = \sum_{i=1}^{39} P_{ei} = P_{e1} + P_{e2} + P_{e3} + \cdots + P_{e39}$$

$$TQ_e = SQ_e/3 \times PQ_e$$

- Quality Assessment Rate in Use

Similarly, to perform a quality assessment in use, the simulator evaluation team assembles an equation with the criteria it considers relevant or can evaluate (Figure 5).

Figure 5: Quality rate in use (TQu). Source: The authors.

$$SQ_u = \sum_{i=1}^{11} P_{ui}S_{ui} = P_{u1}S_{u1} + P_{u2}S_2 + P_{u3}S_{u3} + \cdots + P_{u11}S_{u11}$$

$$PQ_u = \sum_{i=1}^{11} P_{ui} = P_{u1} + P_{u2} + P_{u3} + \cdots + P_{u11}$$

$$TQ_u = SQu/3 \times PQ_u$$

- General Quality Rate (TQg)

The overall quality rate (QG) is the weighted rate of all evaluations performed. If any evaluation has not been performed, its factors are zeroed in the equations presented in Figure 6.

Figure 6: General quality rate (TQg). Source: The authors.

Are: $SQ_g = SQ_e + SQ_u$ and $PQ_g = PQ_e + PQ_u$
The general Quality Rate will be: $TQ_g = SQ_g/3 \times PQ_g$

- Completeness of Evaluations

The evaluation's completeness rate is determined by the number of items evaluated (*n*) on the total of evaluable items. Internal and external quality assessments have 34 items available, while quality assessments in-use have 12. The completeness rates of the evaluations are expressed in Figure 7.

Figure 7: Completeness rates of evaluations. Source: The authors.

Be:
Nea the total of external quality items evaluated.
Nep = 39 the total of external quality items evaluable.
The completeness rate of an external quality assessment is calculated by **TCe = Nea/Nep**
Be:
Nua the total quality items in use evaluated.
Nup = 11 the total quality items in-use evaluable.
The completeness rate of a quality assessment in use is calculated by **TCu = Nua/Nup**
Be:
Nga = (Nea+Nua) the overall total of items evaluated.
Ngp = (Nep+Nup) = 50 the overall total of evaluable items.
The overall completeness rate is calculated by **TCg = Nga/Ngp**

- Sufficiency Factors for Assessment and Functional Adequacy

An evaluation is **considered** sufficient if all functional adequacy items are evaluated, since to make a pedagogical recommendation, it is necessary to evaluate whether the simulator has a functional capacity to assist in achieving desired educational objectives. Otherwise, the evaluation is considered **insufficient.** A simulator is considered **functionally adequate** if all adequacy items are evaluated with a score equal to 2 (satisfactory with restrictions) or equal to 3 (satisfactory without restrictions).

- Star Rating

If the evaluation is sufficient, the quality of the simulator can be classified on a scale from zero to five stars (Figure 8), in accordance with the functional adequacy and the result of the General Quality Assessment Rate **(TQg)**. It is important to highlight that, for a course coordinator or a teacher – decision-makers regarding the adoption or not of a particular simulator in a course or class – it is important to know not only the quality classification calculated for the simulator but also the rate of completeness of the evaluation performed. The higher the completeness rate, the more comprehensive the evaluation performed.

Figure 8: Calculation of star rating (pseudocode). Source: The authors.

```
If insufficient evaluation
Then Stars = unrated
Else If simulator not functionally suitable
   Then Stars = 0
   Else Stars = TQg / 20 (round to a decimal place)
```

Evaluation of the External Quality of Physics Simulators

Before the evaluation of the simulators themselves, the teachers made comments about the evaluation model. Despite the small number of evaluating teachers in the model (2 teachers), it was possible to verify that the result was satisfactory, as it is a consensus among the teachers that MAPHYSE meets the expectations for the evaluation of simulators/virtual laboratories. Table 13 displays the teachers' comments on the proposed simulator evaluation model for this research.

Table 13: General comments from the teachers about MAPHYSE

Teacher	Comment
A	Very interesting. I had never evaluated a simulator before. It's good to have a model for simulator evaluation. It meets the objectives of assessing a simulator.
B	The MAPHYSE model allows for the evaluation of virtual laboratories and, consequently, provides feedback to the teacher regarding their quality.

The evaluations of physics simulators presented in this section were performed through evaluation processes of external quality and quality in-use aspects, with an interdisciplinary team composed of a physics teacher and one of Computer Science, both with 15 years of teaching experience (Figure 9).

Figure 9: MAPHYSE parameters of quality assessments. Source: The authors.

pLang = Brazilian Portuguese; **qMe** = 100 runs; **pTimeA** = 3s; **pTimeB** = 5s;
pSpaceA = 1MB; **pSpaceB** = 2MB

Tables A1 and A2 (Appendix A) list, respectively, the external quality sub-characteristics selected and those not selected in this evaluation. Regarding the quality in use, tables A3 and A4 (Appendix A) list, respectively, the selected and unselected sub-characteristics in this evaluation. Considering the items with weight greater than zero, this evaluation process' completeness rates are presented in Figure 10.

Figure 10: Calculations of completeness rates. Source: The authors.

Nea=18, **Nep**=39, **Nua**=4, **Nup**=11, **Nga=Nea+Nua**=22, **Ngp=Nep+Nup**=50
The completeness rate of:
- external quality is calculated by **TCe** = Nea/Nep = 18/39 = **46.2**
- quality in use is calculated by **TCu** = Nua/Nup = 4/11 = **36.4**
- overall quality is calculated by **TCg** = Nga/Ngp = 22/50 = **44.0**

- The *Physic Virtual Lab* Application

For the present study, three simulators of the educational mobile application *Physic Virtual Lab* (PVL)[9], which is available for the Android platform, were analyzed. This application was selected for its completeness: a suite composed of 67 simulations of physical phenomena and experiments. In this application, the simulations are distributed in the following areas: mirrors and lenses, mechanics, electricity and magnetism, waves and thermodynamics. The student can change parameters or move objects and perceive the results. Equations of the experiments are provided for better understanding. Being all simulators belonging to the same application, Table A5 and A6 (Appendix A) present, respectively, common sub-characteristics of external quality and in use.

- *Net Vector* Simulator

By definition, a vector is a mathematical entity represented by a straight oriented segment. The length of this line segment represents the vector module; the straight support of the segment determines its direction. The orientation of the line segment indicates the direction that should be represented by employing an arrow (Souza, 2015).

[9] Installable via Google Play Store.

The *Net Vector simulator* shown in Figure 11 (left) is recommended for mechanics' studies. It is proposed to illustrate the representation of a vector with its components in the four quadrants. There are two sliding buttons to change the green and blue vector sizes and two more to change their angles (*green and blue vector angles*). The vector represented in Figure 11 (right) has oblique direction and direction from left to right, bottom-up.

The evaluation reports of the sub-characteristics of external quality and quality in-use of this simulator are presented, respectively, in tables A7 and A8 (Appendix A). Figure 12 presents the calculations of quality rates.

Figure 11: Net Vector Simulator (left). Source: PVL. A vector (right). Source: Souza (2015)

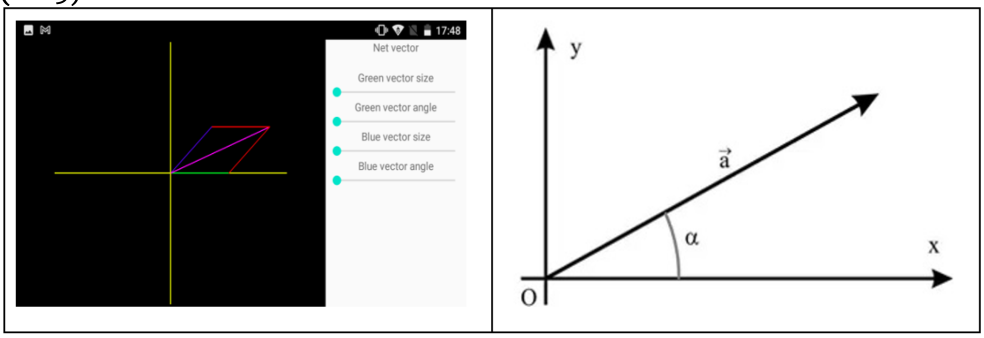


Figure 12: Net Vector simulator Quality Rates. Source: The authors.

SQe = 73; **PQe** = 38; **TQe** = 64.0; **SQu** = 17; **PQu** = 9; **TQu** = 63.0;
SQg = 90; **PQg** = 47; **TQg** = 63.8
Stars = 0

The accuracy problems observed in the sub-reporting of Table A7 (Appendix A) are severe quality problems that have the potential to reinforce alternative concepts instead of scientific ones. As in the simulator, vector projections are found in space. A significant risk in the learning context is that students can confuse the projections of the vectors that should be on the Cartesian Orthogonal System's axes and not in space. Another problem is related to the vector concept itself concerning its direction, module, and angle, which are not clear in the simulator. According to the functional adequacy criterion, the *Net Vector* simulator has, therefore, zero stars, and it is not recommended to be used in schools in their teaching and learning processes. In the other quality sub-characteristics evaluated, there is also no accessibility treatment.

- *Electric Field Vectors* Simulator for Negative and Positive Source Charges

Another experiment in Electricity and Magnetism is the visualization of electrical charges, which cannot be carried out through a real experiment. One can only visualize the effect of the behavior of electrical charges. To better understand the electric field, we use some models of representation of the force lines that are imaginary lines drawn so that their tangent, at any point, points in the direction of the vector of the electric field at that point. The proximity between them is related to the intensity of the electric field in that region of space. The closer the electric field lines are, the more intense the electric field in that region and vice versa.

The *Electric Field Vectors* for a Negative and a Positive Source Charges (EFVNPSC) simulator, whose screenshot is shown by Figure 13 (left side), uses the imaginary line model to represent the electric field. It has two electrical charges: one negative (left) and one positive (on the right). Among them, the electric field lines are represented. On the right are buttons that handle the intensity of negative and positive charges and the separation distance between them, thus varying the electric field representation lines' configuration.

By convention, the power or field lines (imaginary) have the same meaning of the electric field vectors and have as characteristics the following properties (Figure 13, right side): a) come out of positive loads and arrive at negative loads; b) the force line is tangent to the electric field vector at each of its points; c) by one point of an electric field may not pass two power lines. Therefore, two power lines cannot cross, and d) the electric field's intensity is proportional to the power lines' concentration.

The evaluation reports of the sub-characteristics of external quality and quality in-use of this simulator are presented, respectively, in tables A9 and A10 (Appendix A). Figure 14 presents the calculations of quality rates.

e-ISSN: 2595-4881

Figure13: Electric field line simulator (left). Source: PVL. Electric field lines for two distinct loads (right). Source: Moraes *et al.* (2019)

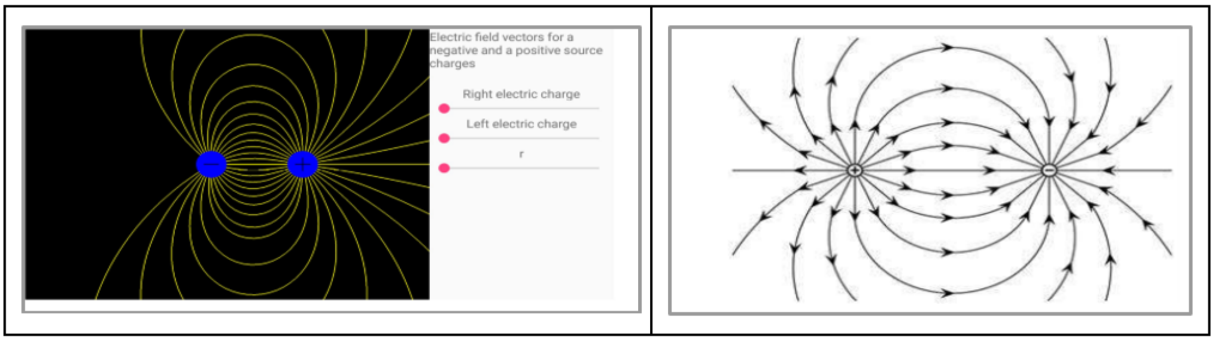


Figure 14: EFVNPSC Quality Rate. Source: The authors.

SQe = 83; **PQe** = 38; **TQe** = 72.8; **SQu** = 20; **PQu** = 9; **TQu** = 74.1;
SQg = 103; **PQg** = 47; **TQg** = 73.0
Stars = 3.7

This simulator was evaluated as suitable with restrictions, rated at 3.7 stars. Despite the accuracy and capacity problems observed, EFVNPSC may be recommended for pedagogical activities. Whenever using this simulator in classes, it is recommended that the teacher draw attention to the power lines' characteristics. Only then, the student can better understand how the force of interaction between the loads in terms of direction is and will also better understand the orientation of the electric field in any region of the space where the effect of this is felt. On the other hand, the student must understand that, for equal loads, the number of electric field lines must always be the same for each of the loads. There are also opportunities for improvements in accessibility.

Final Remarks

The evaluations carried out based on MAPHYSE allow us to affirm that the ISO/IEC 25010 standard supports the creation of very complete reference models for external quality assessment and in-use. Also, it is feasible to select quality items to be evaluated, provided that functional adequacy is included since software must offer the functionalities to which it is proposed.

Interdisciplinary teams can use MAPHYSE to evaluate various Physics simulators. Since these evaluations are carried out by the same criteria and, consequently, with the same completeness, it allows educators to make decisions that are more based, judicious of adopting simulators in the educational processes in which they operate. Besides, it allows simulator developers to *receive qualified feedback* scans that provide them with improvements and fixes.

Concerning the simulators evaluated, it is concluded that *Net Vector* is not recommended for didactic purposes since it can reinforce alternative conceptions about the content worked. Although susceptible to recommendation, the other also presents opportunities for improvement, especially in relation to functional adequacy, capacity, and accessibility.

As future works, two sets of action are being planned. One of them is to evaluate more simulators based on MAPHYSE. The other is to increase the completeness of evaluating simulators already evaluated and recommended (i.e., those with three or more stars) with research present in classrooms.

Finally, it is understood that just as the ISO/IEC 25010 standard enabled the creation of MAPHYSE, it is possible to create reference models for evaluation of internal, external, and in-use quality for evaluations of other categories of software based on this Standard.

Referências

BASTOS, L.; WILKINSON, R. Análise Estatística de Simuladores. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 19, 2010, Florianópolis. **Anais...** Florianópolis: UNICAMP, 2010.

BRITTO, T. C. P. **GAIA:** uma proposta de guia de recomendações de acessibilidade web com foco em aspectos do autismo. 2016. 257 f. Dissertation (Master in Computer Science) – Universidade Federal de São Carlos, São Carlos, 2016.

CAMPÊLO, R. A. **Acessibilidade para portadores de dislexia em um ambiente virtual de aprendizagem móvel**. 2013. 231 f. Dissertation (Master in Computer Science) - Universidade Federal de Pernambuco, Recife, 2013.

CONDE, B. J. M. **Das conceções alternativas às conceções científicas com metodologias ativas de aprendizagem e utilização de simuladores**: uma intervenção didática para a aprendizagem da Física do som. 2021. 152 f. Dissertation (Master's in Social Sciences) - Instituto Politecnico de Leiria Leiria, Portugal, 2021.

DA SILVA, R. S. et al. Avaliação de Software Educativo: a complexidade de escolher uma abordagem adequada. In: CONGRESSO REGIONAL SOBRE TECNOLOGIAS NA EDUCAÇÃO, 1, 2016, Natal. **Anais...** Natal: UFRN, 2016.

DE MENEZES LEÃO, N. M.; KALHIL, J. B. Concepções alternativas e os conceitos científicos: uma contribuição para o ensino de ciências. **Latin-American Journal of Physics Education**, v. 9, n. 4, p. 12, 2015.

DE OLIVEIRA, L. B. R. **Architectural design of service-oriented robotic systems**. 2015. 187 f. Thesis (Doctorate in Computer Science) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2015.

FERNÁNDEZ, J. L. S. **Estudio de la influencia de un entorno de simulación por ordenador en el aprendizaje por investigación de la Física en Bachillerato**. Salamanca: Secretaría General Técnica. Subdirección General de Información y Publicaciones, 2005. 167 p.

FERNÁNDEZ, J. L. S. Informática y enseñanza de las ciencias. In: TEIXEIRA, P. et al. (Org.). **Didáctica de las Ciencias Experimentales**: teoría y práctica de la enseñanza de las ciencias. Barcelona: Paidós, 2000. p. 339-360.

HERPICH, F. et al. Modelo de avaliação de abordagens educacionais em Realidade Aumentada Móvel. **RENOTE**, v. 17, n. 1, p. 355-364, 2019.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION/INTERNATIONAL ELECTROTECHNICAL COMMISSION. **ISO/IEC 25010**: Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models. Genebra: ISO/IEC, 2011.

KOÇAK, S. A.; ALPTEKIN, G. I.; BENER, A. B. Integrating Environmental Sustainability in Software Product Quality. In: INTERNATIONAL WORKSHOP ON REQUIREMENTS ENGINEERING FOR SUSTAINABLE SYSTEMS, 4, 2015, Ottawa. **Proceedings...** Ottawa: [s.n.], 2015. p. 17-24.

MATHIS, F.; VANIEA, K.; KHAMIS, M. Replicueauth: Validating the use of a lab-based virtual reality setup for evaluating authentication systems. In: *CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS*, 7, 2021, Yokorama. **Anais...** Yokorama: [s.n.], 2021. p. 1-18.

MORAES, I. P.; ALVES, R.; DE NOVAIS, E. R. P. Experimento para visualização das linhas de campo elétrico. **Scientia Plena**, v.15, n.7, 2019.

NUNES, M. S. C. **Desafios da inclusão na práxis pedagógica**: saberes e fazeres em ciência da informação. São Paulo: ABECIN Editora, 2020.

PETTICREW, M.; ROBERTS, H. **Systematic reviews in the social sciences**: a practical guide. Hoboken: John Wiley & Sons, 2008.

QUEIROZ, G. R. P. C.; BARBOSA-LIMA, M. D. C. A. Conhecimento científico, seu ensino e aprendizagem: atualidade do construtivismo. **Ciência & Educação** (Bauru), v.13, n.3, p.273-291, 2007.

SCHROEDER, E. Conceitos espontâneos e conceitos científicos: o processo da construção conceitual em Vygotsky. **Atos de pesquisa em educação**, v.2, n.2, p.293-318, 2007.

SILVA, C. L. D. **Desenvolvimento de um software para implantação do processo de enfermagem**. 2015. 114 f. Dissertation (Masters in Nursing) – Universidade do Vale do Rio Sinos, Porto Alegre, 2015.

SIQUEIRA, I. C. **Proposta de modelo de avaliação no Simulador Virtual Tático com base nos simuladores utilizados no Exército Brasileiro**. 2022. 61 f. Monograph (Specialization in Missile and Rocket System Operation) – Centro de Instrução de Artilharia de Mísseis e Foguetes, Formosa, 2022.

SOAD, G. W. **Avaliação de qualidade em aplicativos educacionais móveis**. 2017. 147 f. Dissertação (Mestrado em Ciências) - Universidade de São Paulo, São Carlos, 2017.

SOUZA, N. L. de. Grandezas escalares e vectoriais, 2015. Disponível em: http://educacao.globo.com/fisica/assunto/mecanica/grandezas-escalares-e-vetoriais.html. Acessado em: 15 abr. 2023.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL. Acessibilidade digital: tipo de fonte. UFRGS: CPD, 2020. Disponível em: https://www.ufrgs.br/acessibilidadedigital/tipo-de-fonte.

WORLD HEALTH ORGANIZATION. World report on disability 2011. Geneva: World Health Organization, 2011.

APPENDIX A – Simulator Quality Assessment Tables

Source: The authors.

Table A1: Selected external quality sub-characteristics.

Ref.	Weight	Ref.	Weight	Ref.	Weight	Ref.	Weight
S_{e1}	3	S_{e8}	1	S_{e14}	2	S_{e19}	2
S_{e2}	3	S_{e10}	2	S_{e15}	2	S_{e21}	1
S_{e3}	3	S_{e11}	2	S_{e16}	2	S_{e22}	2
S_{e4}	3	S_{e12}	2	S_{e17}	2	S_{e29}	3
		S_{e13}	2	S_{e18}	1		

Table A2: Unselected external sub-characteristics

Ref.	Non-Selection Justification (Zero Weight)
S_{e5} .. S_{e7}, S_{e20}	The simulators do not have fault tolerance mechanisms, don't supply test roadmap, do not depends on remote functionalities, do not have bandwidth consumption.
S_{e9}, S_{e23} .. S_{e30}, S_{e31} .. S_{e39}	Quality in use, data imports/exports, communication with servers, adaptability, ability to replace and the ability to coexist not scoped.

Table A3: Selected quality sub-characteristics in use

Ref.	Weight	Ref.	Weight	Ref.	Weight	Ref.	Weight
S_{u4}	2	S_{u7}	2	S_{u10}	3	S_{u11}	2

Table A4 Quality in use: unselected sub-characteristics

Ref.	Non-Selection Justifications
S_{u1}, S_{u2}, S_{u4} .. S_{u6}, S_{u8}, S_{u9}	Effectiveness, efficiency in use, utility, pleasure and comfort requirements, evaluation of ergonomics, energy efficiency and network consumption not scoped.

Table A5: Evaluation report of external quality sub-characteristics common to all simulators

Evals.	Observations
$S_{e11} = 1$	The simulators provide only English language.
$S_{e13} = 1$	Only English language and no verbalization about the physical effects, not offering accessibility for the visually impaired.
$S_{e14} = 2$	The simulators do not require hearing capacity. Thus, it meets the needs of deaf people or people with low hearing acuity, unless who have joint needs for visual or auditory accessibility.

$S_{e21} = 3$	The application demanded an average of 0.13 MB per simulator, below the limit of 1MB set **in pEspacoA**.
$S_{e29} = 2$	Available only for Android (not available for iOS).

Table A6: Evaluation report of quality sub-characteristics in use common to all simulators

Evals.	Observations
$S_{u4} = 3$	On January 7th, 2021, the simulators were rated with 4.1 stars at the application store indicating confidence in their ability to assist them in their respective learning activities.
$S_{u7} = 3$	Free app.

Table A7: Net Vector External Quality Assessment Report

Evals.	Observations
$S_{e1} = 1$	It does not project the vector on the axes, not showing sizes and angles variations.
$S_{e2} = 1$	It does not show module and angle values.
$S_{e3} = 1$	It does not achieve the objectives of learning.
$S_{e4} = 3$	In 100 executions of the simulator performed there were no failures.
$S_{e8} = 3$	The key concepts allow the user to understand how to manipulate the simulator with ease.
$S_{e10} = 3$	It is easy to use and operate.
$S_{e12} = 3$	It has a standardized user-friendly interface.
$S_{e15} = 1$	It does not provide alternatives for people with motor difficulties to use sliding buttons.
$S_{e16} = 1$	The color spectrum presents very intense colors, not recommended for autistic people.
$S_{e17} = 3$	The interface is straightforward, with common sliding buttons. The unjustified short texts are adequate for dyslexics.
$S_{e18} = 3$	It prevents students from making operational errors because it is simple to manipulate.
$S_{e19} = 3$	The response time is satisfactory (around 2s for the opening of the simulation) and almost instantaneous when manipulating buttons.
$S_{e22} = 2$	It does not allow you to exercise negative angles or angles greater than 360 degrees.

Table A8: Net Vector Quality Assessment Report

Evals.	Observations
$S_{u10} = 1$	It compromises the objectives of learning.
$S_{u11} = 1$	It does not offer additional characteristics to complement learning.

Table A9: EFVNPSC External Quality Assessment Report

Evals.	Observations
$S_{e1} = 3$	It illustrates the model of representation of electrical forces through imaginary lines.
$S_{e2} = 2$	The direction of the electric field lines is not shown. It also fails to represent the proportionality of the intensity of the electric field with the concentration of the power lines. Therefore, it needs accuracy adjustments.
$S_{e3} = 2$	It allows the objectives of learning to be partially attained.
$S_{e4} = 3$	In 100 executions performed there were no records of failures.
$S_{e8} = 3$	It allows the student easily to understand the key concepts and how to manipulate the simulator.
$S_{e10} = 3$	It is easy to use because its interface is based on drag buttons.
$S_{e12} = 3$	It is pleasant and simple to use and observe the phenomena.
$S_{e15} = 1$	It does not provide alternatives for people with motor difficulties to use sliding buttons.
$S_{e16} = 1$	The color spectrum presents very intense colors, not recommended for autistic people.
$S_{e17} = 3$	The interface is straightforward, with common sliding buttons. The unjustified short texts are adequate for dyslexics.
$S_{e18} = 3$	It prevents students from making operational errors because it is simple to manipulate.
$S_{e19} = 3$	The response time is satisfactory (around 2s for the opening of the simulation) and almost instantaneous when manipulating buttons.
$S_{e22} = 1$	To exercise electric fields, it is necessary to measure positive/negative charges in Coulombs and its submultiples. Rays are measured in meters and its submultiples. The simulator does not inform the units of measure. It therefore is not feasible to evaluate the ability of the simulator.

Table A10: Quality assessment report in use of EFVNPSC

Evals.	Observations
$S_{u10} = 2$	The simulator offers necessary experimentation contexts that allow us to partially achieve educational objectives.
$S_{u11} = 1$	The simulator does not offer other contexts of experimentation other than those necessary for the achievement of the learning objectives.

[4] Addresses: < http://portal.acm.org, www.engineeringvillage.com, https://scholar.google.com.br, http://ieeexplore.ieee.org, http://www.isiknowledge.com, www.sciencedirect.com, www.scopus.com, http://link.springer.com,www-periodicos-capes-gov-br.ezl.periodicos.capes.gov.br >