INSTITUTO
UNIVERSITÁRIO
DE LISBOA

DATUS: Dashboard Assessment Usability Model - a case study with student dashboards

Rita Sofia da Silva Antunes

Master in, Computer Science and Business Management

Supervisor:
Doctor Elsa Alexandra Cabral da Rocha Cardoso, Assistant Professor,
Iscte – Instituto Universitário de Lisboa

Co-Supervisor:
Doctor José Eduardo de Mendonça Tomás Barateiro, Invited Assistant Professor,
Iscte – Instituto Universitário de Lisboa

November, 2020

TECNOLOGIAS
E ARQUITETURA

DATUS: Dashboard Assessment Usability Model - a case study with student dashboards

Rita Sofia da Silva Antunes

Master in, Computer Science and Business Management

Supervisor:
Doctor Elsa Alexandra Cabral da Rocha Cardoso, Assistant Professor,
Iscte – Instituto Universitário de Lisboa

Co-Supervisor:
Doctor José Eduardo de Mendonça Tomás Barateiro, Invited Assistant Professor,
Iscte – Instituto Universitário de Lisboa

November, 2020

Acknowledgements

To my parents who encouraged me to go on this adventure and never let me lack anything.

But specially to my mother, who was the most present person and always picked me up at my lowest moments and when I doubted that I would make it to the end, she always believed for both.

To my sisters who are a source of inspiration and support and were always there if I needed them, even in separate houses.

To my special friends that along the way we shared moments of all kinds and that helped me turn on the motivation button.

To my supervisors for their endless support and effort. To Professor José Barateiro for all the inputs this dissertation needed, and Professor Elsa Cardoso who has been a constant source of knowledge and inspiration, my sincere thank you for their guidance.

Resumo

Português

O mercado de software observa o aparecimento de novas empresas e produtos todos os dias. Este crescimento traduz-se em competição e a sobrevivência das empresas resume-se ao investimento nos seus produtos. Também as universidades têm interesse em melhorar o seu produto, o ensino. Esta melhoria pode ser alcançada através de investimento na experiência de aprendizagem dos estudantes. A usabilidade e a experiência do utilizador desempenham um papel importante e demonstram ser uma vantagem competitiva em que vale a pena investir.

Consequentemente, têm surgido novos métodos para melhorar o processo de avaliação de usabilidade. Apesar deste crescimento, não existe um modelo claro para avaliar a usabilidade de um dashboard. Esta lacuna levou à investigação desta dissertação, uma proposta de um novo modelo, Dashboard Assessment Usability Model (DATUS), acompanhado por um método de avaliação, que pode ser aplicado à avaliação da usabilidade de dashboards.

Estão incluídas no DATUS oito dimensões de usabilidade, cada uma corresponde a uma faceta específica de usabilidade que foi identificada numa normalização ou modelo existente, e decompõem-se num total de 20 métricas.

Para verificar se o modelo é viável, e como contribuição para o Iscte - Instituto Universitário de Lisboa, foi desenhado um protótipo de dashboard para a plataforma Fénix, à qual o modelo DATUS foi aplicado.

Para testar a usabilidade dos dashboards, foi realizado um estudo comportamental com 30 alunos do Iscte. Após a análise dos resultados, foi confirmada a viabilidade do modelo e do método propostos e retiraram-se conclusões positivas em relação à usabilidade do protótipo.

Palavras-chave: Business Intelligence, usabilidade, dashboard, questionário, experiência de aprendizagem

Abstract

English

The software market sees the appearance of new companies and products every day. This growth translates into the competition, and the survival of companies is reduced to investment in their products. Universities are also interested in improving their product, education. This improvement can be achieved by investing in the learning experience of students. Usability and user experience play an important role and have been a competitive advantage worth investing.

Consequently, new methods have emerged to improve the process of evaluating the usability of products. Despite this growth, there is no direct model for assessing the usability of a dashboard. This gap led to the investigation of this dissertation, a proposal for a new model, Dashboard Assessment Usability Model (DATUS), accompanied by an evaluation method, which can be applied to the evaluation of the usability of dashboards.

Eight usability dimensions are included in DATUS, each corresponding to a specific usability facet that has been identified in an existing standard or model and decomposed into a total of 20 metrics.

In this sense, to verify if the model created is feasible, and as a contribution to Iscte - Instituto Universitário de Lisboa, a prototype dashboard was designed for the Fénix platform, to which the DATUS model was applied.

To test the usability of the dashboards, a behavioural study was conducted with 30 Iscte students. After analysing the results, not only was the feasibility of the proposed model and method confirmed, but positive conclusions were also reached regarding the usability of the prototype.

Keywords: Business Intelligence, usability, dashboard, questionnaire, learning experience

Index

Index of contents and figures

Contents

Figures

Glossary of acronyms

AA	Academic Analytics
AMBA	Analyse my Blackboard Activities
BI	Business Intelligence
BIA	Business Intelligence and Analytics
CSUQ	Computer System Usability Questionnaire
DATUS	Dashboard Assessment Usability Model
DSR	Design Science Research
DSS	Decision Support Systems
HCI	Human-Computer Interaction
IEC	International Electrotechnical Commission
Iscte	Instituto Superior das Ciências do Trabalho e da Empresa
ISO	International Standards Organization
KPI	Key Performance Indicators
LA	Learning Analytics
LAD	Learning Analytics Dashboard
OLAP	Online Analytical Processing
PSSUQ	Post-Study System Usability Questionnaire
SUMI	Software Usability Measurement Inventory
SUS	Software Usability System
QUIS	Questionnaire for User Interaction Satisfaction
UMUX	Usability Metric for User Experience
USE	Usefulness, Satisfaction, and Ease of use
UX	User Experience

Introduction

"To achieve great things, two things are needed: a plan, and not quite enough time."

(Leonard Bernstein)

Usability is present in all the products we use in our daily life [1], from the simplest kettle to heat the morning coffee to something more complex like the interaction with a dashboard that monitors the vital signs of patients in a hospital. Logically misinterpretation or use of a toaster can leave the breakfast unfit for consumption, but when we are dealing with the example of the dashboard in hospitals, the seriousness grows exponentially. In these situations, the problem is not in the user but the poor usability that a product can have.

The significative growth of the software market increases the production of digital products and causes new companies and new products to be created every day, leading to a constant battle between companies in the market in the pursuit of competitive advantages, that is, means to stand out from the competition. However, quantity does not always mean or is accompanied by quality. A product or system can be designed containing the finest functionalities, but what indeed attaches users to the product rely on the interaction, and the more intuitive this interaction is, the better. It is in this field that the inferences about usability and the concepts around it have a complementary but fundamental role.

Not selling products and be defeated by competition is not limited to lack of quality but is closely linked to usability. Measuring usability is an expensive and time-consuming process [2], and many companies fail to invest sufficient resources in this field, causing the loss or even the never arrival of new consumers/users. Investing in usability is investing in a product to be usable. Even if there is an investment from companies, the effort to measure usability and establish formal metrics could be wasted without a consolidated framework [3].

The process of measuring usability relies on collecting metrics to measure users' performance [4]. Moreover, since usability and user experience have been a worthy competitive advantage to invest in, several methods have emerged [5][6][7]. Despite this growing number of studies, when it comes to interactive graphical interfaces, such as dashboards, there is no straightforward method or model to assess its usability. This gap led to the research carried out in this dissertation, with a proposal of a new model, accompanied by an evaluation method, that can be applied to the usability assessment of dashboards.

Although academic organisations do not emerge every day as companies, universities seek ways to improve their product: education. Improving education can mean improving the learning experience and creating greater engagement with students. Perhaps understanding these concepts is essential before going further. The learner experience refers to any interaction in which learning takes place, and the learner engagement refers to the portion of physical and psychological energy dedicated by students to the learning experience [8]. It is important to note that students' engagement is related to academic

achievement, satisfaction, and sense of community [9]. These latter are aspects in which universities can invest to improve their product.

Education in universities has the most independent character among educational institutions. This is reflected in greater responsibility for students in their own academic achievement and higher interest in keeping track of their performance. Universities store ample amounts of student data such as entry qualifications or daily interaction with learning platforms, i.e., data regarding the student learning experience. By aligning the new field of learning analytics, universities can provide *actionable intelligence* about individuals and groups of students [10]. There is thus an opportunity for universities to invest in the learning experience of their students, and hopefully increase student's engagement, by providing them with a tool in which they can monitor their performance and then make decisions with a view to their individual academic goals, such as a student dashboard.

The importance of empowering students with an analytical dashboard has been sustained with several studies, mentioned in the following chapter. Reviews of this kind do change the way students monitor their learning experience and control their performance.

The Learning Analytics and Educational Technology research community have been focusing on the study of student dashboards, and several papers have been published in the last three years. Therefore, the study of student dashboards to monitor their learning experience is a quite recent and relevant scientific topic. In order to contribute to research on student dashboards, identifying usability issues and information presentation improvements, the need for a more practical solution for this process emerges.

It is through this interconnection between the importance of usability in product development, and consequently in dashboards, and the intention to look for ways to improve the product offered by educational institutions that the research and the development of two artefacts of this dissertation arises. Following a design science research (DSR), the two artefacts cover a new model and a method with the objective of evaluating the usability of a dashboard independently of its context and presented information. Along with these artefacts, a dashboard will be built to validate and assess the model and method's applicability. The dashboard will be a proposal for Iscte - Instituto Universitário de Lisboa[1] and aims to use the best practices of Business Intelligence to monitor the student learning experience. The dashboard proposal is intended to monitor the students' performance continuously throughout the academic year, displaying relevant information regarding students' performance and learning experience.

[1] From here after Iscte - Instituto Universitário de Lisboa will be referred as Iscte.

Contributions and objectives

This dissertation addresses a recent and relevant topic to contribute to the scientific community, as stated before, and the primary focus of this dissertation's research is expressed in terms of the following research question: *how to measure the usability of dashboards, quickly and effectively?*

The expected contributions of this dissertation include the development of a new usability assessment model and method to help dashboard designers to identify usability problems and thus improve their interfaces. To assess the utility of these artefacts it is necessary to apply them to a dashboard. To this end, i.e., to understand if the developed model and method are fully applicable and allow a quick and effective usability evaluation, a proposal of a dashboard for the Fénix platform of Iscte will be designed. This proposal aims to contribute to a future implementation of a tool for students to monitor their performance throughout the academic career.

Methodology

The investigation of this dissertation aims to achieve the study of the applicability of a dashboard evaluation model and the usability of a proposed student dashboard following a design science research methodology. The DSR approach focuses on the design of artefacts with the intention of gathering knowledge and understanding of a problem area in order to improve it [11], that is translated into two objectives: (1) to develop an artefact to solve a practical problem in a specific context and (2) generate new technical and scientific knowledge.

The DSR model, according to [12], consists of six activities in a nominal sequence. The following **Figure 1** demonstrates the research process, which relies on three phases gathering the six activities of the DSR model in order to achieve the objectives previously described.

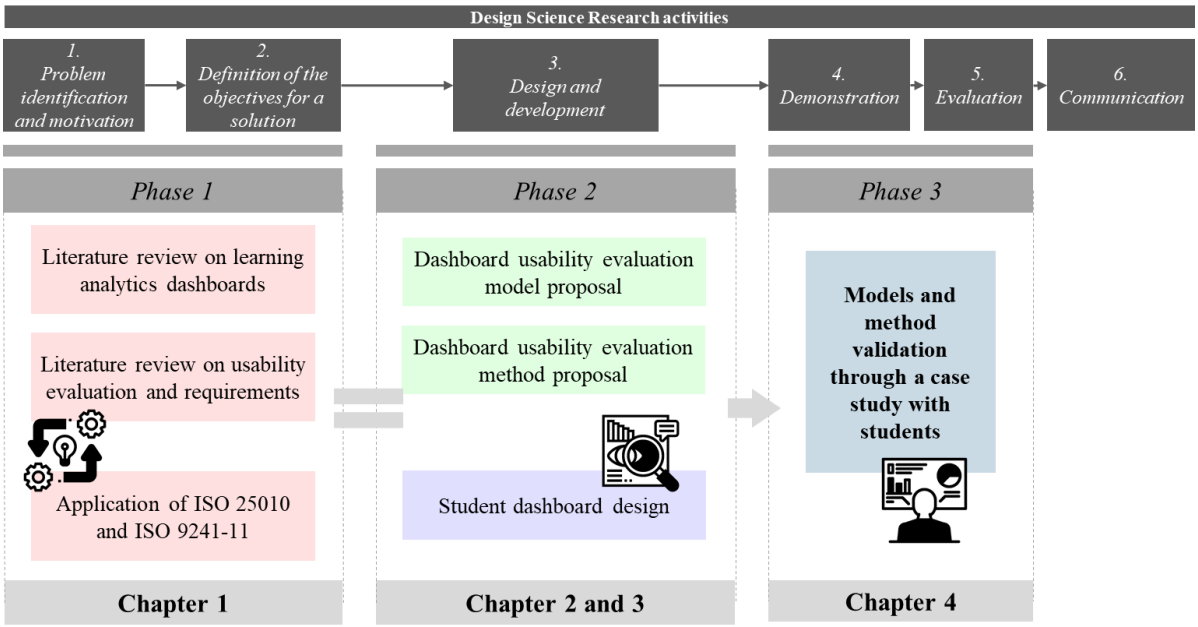

Figure 1: The research design

Phase 1

The initial phase comprises the first two activities of the DSR methodology, that is, the definition of the specific research problem and the justification of the value of the solution to be developed aligned with the knowledge of what is possible and feasible [12].

Phase 1 includes the literature review study focusing on Business Intelligence interfaces, specifically learning analytics dashboards; and usability evaluation and requirements (including models and methods, and the application of usability-related standards of the International Organization for Standardisation (ISO)).

Phase 2

Following the next activities of the DSR model, this phase comprises the development of the artefacts and the demonstration of its use to solve the instances of the problem previously defined.

Phase 2 relates to the main developments and research contributions of the dissertation. The design of two artefacts are described: a new model (DATUS) and a method to evaluate the usability of dashboards. Additionally, a student dashboard proposal for Iscte is also described.

Phase 3

The last phase gathers the last three activities of the methodology in one chapter, that is, the use of the artefacts to solve the instances of the problem defined (*demonstration),* the observation and measurement of how well the artefact supports the solution for the problem (*evaluation*) [12].

Phase 3 focuses on the validation of the artefacts produced in the previous phase. The validation includes the assessment of usability aspects of the student dashboard proposal through a behavioural study that uses the defined method and DATUS model. This study will enable the consolidation of the two artefacts, and a validation of their applicability and feasibility to assess the usability of a given dashboard.

The last DSR activity *(communication)* communicates the problem and its importance, the artefacts' utility and effectiveness and comprises the whole dissertation and a future paper regarding the artefacts to be developed.

Structure

The dissertation document is structured as follows: Chapter 1 presents the literature review, corresponding to Phase 1 of the methodology (see **Figure 1**). Chapter 2 details the design of the DATUS model and the proposed method for assessing the usability of dashboards. Chapter 3 describes the design of a student dashboard for Iscte. This design proposal is aimed at undergraduate students and is aligned with the layout of the university' student academic system (i.e., the Fénix platform). Chapters 2 and 3 correspond to Phase 2 of the research methodology. Chapter 4 describes the validation study implemented, corresponding to Phase 3 of the research design. Finally, the closing chapter presents the conclusions, with a discussion of the main contributions and limitations of this dissertation.

Chapter 1:

State of the Art

The matters related to the subject of this dissertation can be found throughout this chapter regarding the literature review. Related work and definitions of concepts are depicted in order to comprehend better what will be forwardly studied and approached.

1.1. Business Intelligence and Analytics

Business intelligence (BI) is a data driven decision support system (DSS) and is the result of a sequence of innovations in the course of time. Annually the number of publications and scientific organisations regarding the adoption of BI techniques, methodologies and tools increases, showing how the domain of BI is rapidly expanding [13] and receiving widespread interest in both industry and academia [14].

In 1989, Howard Dresner used the term Business Intelligence to describe "*concepts and methodologies for the improvement of business decisions using facts and information from supporting systems"* [15]. Expectedly, many authors have been defining BI over the years. From the appraisal of its definitions, the purpose of BI is to give improved support on decision making to create value. BI analyses data from different sources, converts it into useful information and knowledge, through rational analysis [16], and delivers a sort of view, pattern and useful interface [17]. Nevertheless, it is not enough. It is necessary to act. Organisations should look into decision processes to deliver useful information to the decision-makers as intelligence is only produced through making decisions [16].

To [18], analytics is the input to human and automated decision making and answer questions is its job, in simple or complex forms just as [19] states the role of analytics as driving managerial decisions and actions.

Thus, BI and analytics (BIA) are about the development of technologies, techniques, systems, practices, methodologies and applications to analyse critical business data so as to gain new understandings about business and markets and make timely decisions [20].

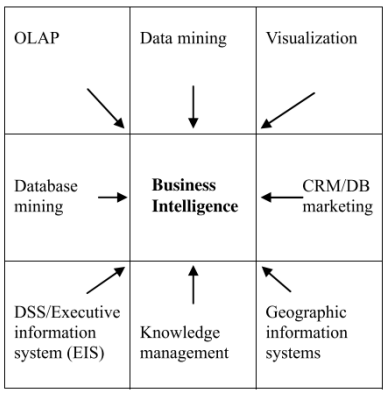

Figure 2: The components of BI [18].

Business intelligence brings multiple components together, as shown in **Figure 2**. From this figure and BI involving many of the analytic techniques of DSSs and in consonance with authors Lim and Chen [21] declared that business intelligence depends upon data analytics, and it is about time to adopt business intelligence and analytics as to the preferred combined term.

From the possible technologies used to achieve its purpose, database query, online analytical processing (OLAP), and advanced reporting tools are generally adopted to explore important data characteristics and scorecards and dashboards to analyse and visualise that data [22]. A brief notion of the concept of these BI interfaces, OLAP and Reports, is presented next.

"*OLAP is the technology behind many BI applications"* [2]. OLAP cubes are the representation of data in matrixes and arrays. It enables looking at information from different perspectives and dimensions and keeping a proper and efficient data structure at the same time [23].

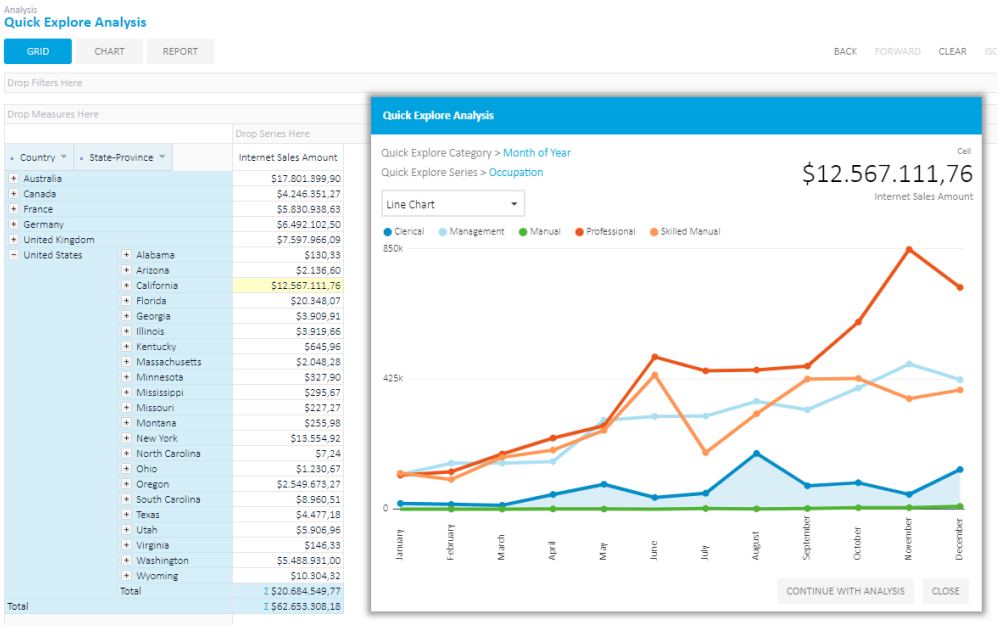

Figure 3: Example of OLAP data analysis in *Kyubit BI platform*

(Source: https://www.kyubit.com/OLAP-Analysis)

In **Figure 2**, reports belong to the "Visualisation" component. Reports enable the data presentation with different formats: graphs, dashboards and performance scorecards. Reports must be endowed with immediate and intuitive interpretation considering the information to be visualised. A Scorecard is a report that measures performance based on objectives, and it is supposed to allow quick reading of the progress towards those objectives, consisting of graphic indicators [24]. Dashboards aggregate several types of reports, and its definition can be found later under the topic of student dashboards.

[2] https://olap.com/

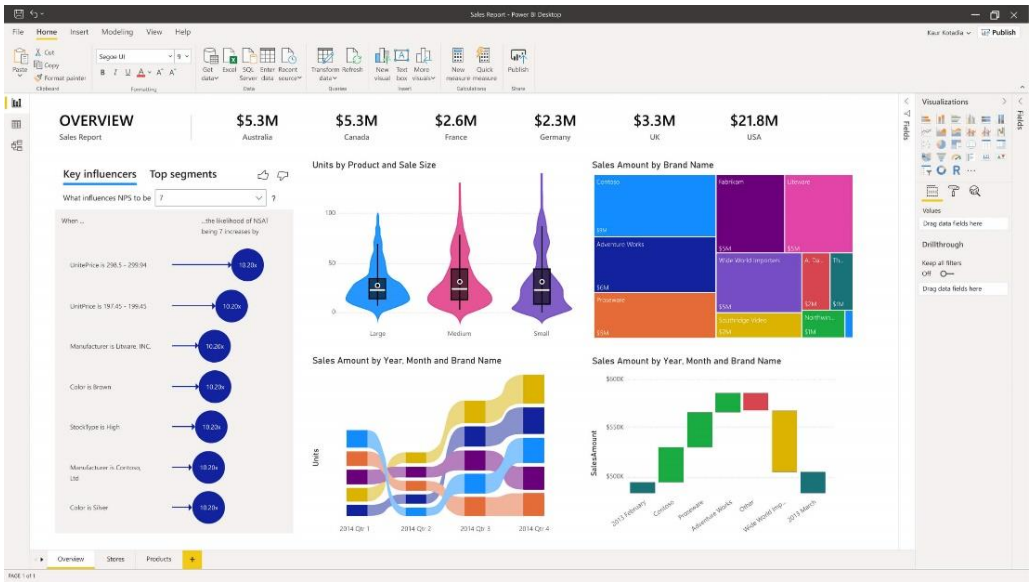

Figure 4: Example of a report in Microsoft PowerBI

(Source: https://powerbi.microsoft.com/)

1.1.1. Learning Analytics

Indubitably, thinking about education in the future highlights new technologies, big data, and analytics. Analytics is present in many fields, such as education. It provides a new model for college and university leaders to improve teaching, learning, organisational efficiency and decision making [25].

Quoting [26], *learning is a product of interaction*. Moreover, from interaction, data can be collected and stored, trends can be identified, and hypotheses can be formed. Analysing data collected from the interaction of users with educational and information technology had attracted considerable attention as an intelligent approach for advancing our understanding of the learning process [21].

In 2011, took place the 1st International Conference on Learning Analytics and Knowledge[3] where Learning Analytics (LA) was defined as the "*measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs"*. This emphasises the potential of LA in positively influencing students' learning and development by enabling higher education institutions to growth their understanding of their students' learning needs [21] as well as help answer educators' questions that reliably arises regarding the performance of the educators' learning and evaluation methods [26].

1.1.2. Academic Analytics

Learning Analytics and Academic Analytics (AA) are two different concepts. In distinction from LA's previous definition, AA is the application of business intelligence in education [27] and highlights analytics at institutional, regional, and international levels [21]. Some authors [10], restrict the definition

[3] https://tekri.athabascau.ca/analytics/

of AA in student retention and graduation rates and declare that AA *"has the potential to create actionable intelligence to improve teaching, learning and student success*". From literature review, frequently the examples discussing AA refers to the problem of perceiving students at risk of dropping out of a course or abandon studies [28]. By that, AA has the potential to improve student retention, enrolment and fundraising [27].

1.1.3. Learning Analytics and Academic Analytics comparison

Although for some years now, according to [29], Academic Analytics has been limited to applying analytical methods to encounter the needs of educational institutions, i.e. in managing enrolment and forecasting academic success through statistics. AA is usually the outcome related to student's progress toward a degree [10]. LA, in turn, besides meeting the needs of educational institutions, its applications can be targeted with particular interest to students and teachers. Thus, LA is increasingly used to achieve goals more aligned with the learning process (i.e. reflection, adaptation, customisation and recommendation). In addition to statistical data such as AA, LA has newer approaches that apply other methods of analysis to improve teaching and performance support.

The following table shows the assessment of these two concepts in the analysis.

Table 1: Learning analytics and Academic analytics [25].

TYPE	LEVEL OF ANALYSIS	WHO BENEFITS
Learning Analytics	**Course-level**: social networks, conceptual development, discourse analysis, "intelligent curriculum."	Learners, faculty
	Departmental: predictive modelling, patterns of success/failure	Learners, faculty
Academic Analytics	**Institutional**: learner profiles, the performance of academics, knowledge flow	Administrators, funders, marketing
	Regional: comparisons between systems	Funders, administrators
	National and International	Nacional governments, education authorities

From the previous statements and the information in **Table 1**, we can gather the following conclusions [21]:

- The focus of LA is on the learning process (analysing the relationship between learners and faculty); therefore, it is more specific than AA;
- AA centres the data analysis at an institutional level, while LA centres on the learning process which allows a better learning experience to the learners and improves the educational quality;
- AA is distinguished similarly to BI, as it raises the need for a model or stage of LA development;

LA is in continuous growth, and the main points that drive the development of this field are the availability of big datasets, the emergence of online learning on a large scale, and political concerns about educational standards [29].

1.2. Student Dashboards

A dashboard is "*a specific type of page where you can insert reports, graphs, charts, and KPI* [4] *lists in order to create a central location for functionally-relevant information"* [30]. It uses simple visual displays (charts, and tables) to communicate BI results, and it can be individually customised according to the user's needs [18]. Not only the design of the interface is a foundation of dashboards but also the accurate capture and selection of the best indicators according to the field. Some authors [31] state the two capabilities of dashboards: the visual power and the way it integrates the primary information a user should remember into one screen regardless of complexity. Dashboard design requires collaboration between the stakeholders and IT specialists [18].

In the educational background, a student dashboard is defined as "*an interactive, historical personalised, and an analytical monitoring display that reflects students' learning patterns, status, performance and interactions"* [32]. It has been created to help students be aware of their performance and engaging with their studies, from a self-assessment perspective and a comparative assessment with their colleagues.

From work found, the following table summarises the currently available information regarding the monitoring indicators previously used to evaluate the usability on student dashboards.

Table 2: Usability metrics identified in previous work

Study	Indicators
Aljohani, N. R., et al., [33] University of Jeddah, King Abdulaziz University, Saudi Arabia	Frequency of accessing the Blackboard Frequency of accessing the discussion board Number of threads added to the discussion board Quiz results
J. Kim, I. H. Jo, and Y. Park [34] Seoul National University, Korea	(Online behaviour patterns from an online course) Login time (total login time for a week) Login frequency (sum of times each student opened LA dashboard) Login regularity (visits on the board and repository) Multiple choice quiz results Final exam grade
Verbert, Duval, Klerkx, Govaerts, & Santos [35] KULeuven, Belgium	Time spent Social interaction Document and tool use Artefacts produced Exercise results/quizzes

[4] Key performance indicators

2.1.1. Learning Analytics Dashboards: student experience

Learning analytics dashboards (LADs) have been deployed in order to help human decision-making by combining two techniques: educational data mining and information visualisation [36]. Which is explained by [37] by showing students' online behaviour patterns in a virtual learning environment. LADs can be used to [38]:

- Provide timely feedback on learning analytics
- Detect students at risk and reducing dropout
- Improve (human) learning
- Improve engagement and motivation
- Support reflection and decision making
- Enhance student's self-awareness and academic achievement

There are several recent studies regarding the applications of dashboards within learning analytics. However, no studies were found regarding academic analytics student dashboards.

According to [32], the authors explored student perceptions and preferences for dashboard features and findings of insights into the attitudes of higher education student groups towards learning through analytics dashboards. They come to conclusions about the content that most interest students to be:

- alerts
- comparisons to peers
- features that support learning opportunities
- preference on having a dashboard of their own to interact and customise/modify the features in their own way

Besides this, a summarised table with the features that students give more importance to is presented. Their findings reveal the potential for providing students with some level of control over learning analytics as a means of increasing self-regulated learning and academic achievement. There is still a discussion between dashboard features being fully customisable and the disagreement of opinions that learning opportunities should be the same for all students.

A study [39] about indicators to design a dashboard was performed with a sample of ten professors. Conclusions showed professors' interest in knowing how students use their course resources and thus could help in the possible redesign of their courses. Indicators that teachers consider the most important include:

- the number of times students accessed the course forum
- the number of forum contributions
- the number of times students consulted the course guide
- number of times each student accessed the course in one week

The conclusions of this study focus the importance of teacher and student training in the pedagogical use of data visualisation.

The authors [33] proposed a framework called "Course-adapted student learning analytics framework" – AMBA[5] tool to empower students with their own analytical dashboard and add the value of the customised student-centred dashboard to the LA application. The experiment leaned on the participation of 86 students in an academic semester. The study was conducted within two groups:

1. Control group, with the teacher-centred dashboard (Blackboard system), where they were told by the teachers about their performance
2. Experimental group, with the student-centred dashboard (AMBA tool), where they could learn about their performance

Conclusions from this study give sustenance to the implementation of a student-centred learning analytics dashboard as:

- The access to the Blackboard system happened more often
- The access to the discussion board happened more often
- Achieved higher final marks
- Improved their engagement with the course more than the results from the control group facilitated the discussion of their own results with their colleagues

An experimental study [34] to evaluate the "usefulness" and "usability" of LA dashboard. The experiment counted with 51 college students at a private university taking an online course, divided into two groups: the experimental group with mobile devices with an awareness reminder mechanism and a control group with regular access to the online course.

The authors collected data from the Learning Management System to measure usage frequency, survey to measure students' satisfaction, and the final scores to measure their learning achievements. Each activity (virtual attendance, individual tasks, quiz, midterm exam, and final exam) contributed to their grades.

Conclusions outlined:

- the students who received dashboard treatment presented higher final score (learning achievement) than those who did not
- a positive correlation between satisfaction with LAD and learning achievement was observed
- the satisfaction from the students who access the LAD more often was lower than those who accessed fewer times. Which indicates a dashboard that piqued interest at first but insufficient to maintain the students' interest.

[5] Analyse my Blackboard Activities

An important finding is that students whose dashboard usage was higher showed lower satisfaction compared to those who used the dashboard fewer times. These students were high academic achievers and tended to be strongly motivated already, so the information from the dashboard could not be required. By this and quoting the authors *"future research needs to consider learners' achievement and motivation level in conducting dashboard treatment"*.

1.3. Principles of Interface Design

Dashboards are data screens designed to help the monitorisation of relevant information. Dashboards must be well designed, i.e. present a valuable interface, in order to guarantee their efficiency. Good design reflects benefits such as communication more effective and less mistaking guidance for users [40]. Principles, rules and heuristics have been presented in the past years in order to help designers deliver interfaces with the maximum usability, even though they happen to overlap each other [41]. Usability of interface design has become particularly crucial because user interfaces should attend to different kinds of needs of different users [40]. Thus, we must underline the authors whose work constitutes baselines to maximise usability in the products or systems of their own designers.

1.3.1. Norman Design Principles

Bad design can bring thoughtful problems. When having trouble with understanding how something is done, it is not the technology that is failing; indeed, the design is [1]. Don Norman, in 1998, identified six fundamental principles of designing that have been key to design digital products and widely used by the Human-Computer Interaction (HCI) community for user interface evaluation and interaction design improvement purposes. The principles are as follows:

- **Visibility**: the more visible an element is, the more probable users will know about it and how to use it.

  It is through observation that users can understand which options are available and what can be done. Therefore, the most important features must be visually well placed and not arouse doubts on how to interact with them.

- **Feedback**: show what action has been taken and what has been accomplished

  This principle is about the output of the user's action. Every user gets greater comfort knowing what their actions triggered, even being a successful outcome or not than to keep on expectations. Feedback can be given through messages, sounds, animations, vibrations or enhancements [41].

- **Affordance**: attributes that allow the user to understand how to use it

  This is the capacity of the object to, speechlessly and with low effort, tell us how they work or should be used.

- **Mapping**: having a clear relationship between controls and their impact

Mapping is about enabling the users to trust that their actions are going to correspond to their expectations.

- **Constrains**: limiting the variety of possible interactions in order to make the interface simpler and adequately guide the user to the next action

The purpose of this principle is to prevent the user from interacting with the system or product wrongly.

- **Consistency**: similarity between operations and elements for achieving similar tasks

It helps users to identify and apply previous patterns towards new situations. There are four types of consistency: aesthetics (appearance), functional (meaning and action), internal (elements and panels) and external (other elements).

1.3.2. Shneiderman Eight Golden Rules[6]

It is equally important to reference the following rules that are part of Ben Shneiderman's book *Designing the User Interface: Strategies for Effective Human-Computer Interaction* to attend as a guide to robust interaction design:

1. Strive for consistency.
2. Seek universal usability.
3. Offer informative feedback.
4. Design dialogues to yield closure.
5. Prevent errors.
6. Permit easy reversal of actions.
7. Keep users in control.
8. Reduce short-term memory load.

1.3.3. Nielsen Heuristics[7]

Heuristic evaluation is used in HCI studies to measure the usability of information systems [42]. There are ten general principles, called heuristics, that are broad rules and not specific guidelines [43].

1. **Visibility of system status** – how well the state of a system is conveyed to its users. It enables users to understand what is happening and make decisions based on that information.
2. **Match between system and the real world** – the system or product should attend the needs of the different types of users.
3. **User control and freedom** – users, can accidentally press a button or a feature and may need a clear way to back out when needed, giving them confidence exploring and using the product.

[6] https://www.cs.umd.edu/users/ben/goldenrules.html
[7] http://www.useit.com/papers/heuristic/heuristic_list.html

4. **Consistency and standards** – consistency with the elements and their attributes should happen internally and externally.

5. **Error prevention** – as one of the most important heuristics, especially when designing under high stakes that can compromise, for instance, safety.

6. **Recognition rather than recall** – to design thinking on minimising user's effort providing cues in the context

7. **Flexibility and efficiency of use** – facilitate the user's effort by enabling the user to pick whichever fits and is comfortable for them

8. **Aesthetic and minimalist design** – making sure the content and visual design focuses on the essentials

9. **Help users identify and recover from errors** – inform users when an error has occurred, what went wrong and offer a solution

10. **Help and documentation** – make help easy to search, focused on the user's task and use lists as steps to be followed

1.4. Software Quality

The term *quality* seems self-explanatory by itself, but there are many different sights of its meaning and how it should be achieved as part of a software production process [4]. There are two critical meanings highly important to manage quality [44]: *features of product* and *freedom from deficiencies.* Their handbook states these meanings along the lines [44]:

1. Quality consists of products' features which meet the need of customers providing product satisfaction;
2. Quality means freedom from errors that require rework or that result in field failure, customer dissatisfaction, claims and so on.

Quality is a complex and multifaceted concept [45] and software quality has been characterised, by many authors and the International Organization for Standardization, in order to accomplish a complete understanding of the concept. According to ISO/IEC 20510:2011[8], *quality* is guided by the *quality in use model* and the *product quality model*.

Table 3: Definitions by ISO 25010 [46]

ISO/IEC 25010:2011 DEFINITIONS	
Quality in use model	"Degree to which a product or system can be used by specific users to meet their needs to achieve specific goals with effectiveness, efficiency, freedom from risk and satisfaction in specific contexts of use."
Product quality model	"Composed of eight characteristics that relate to static properties of software and dynamic properties of the computer system."

[8] https://iso25000.com/index.php/en/iso-25000-standards/iso-25010

Thusly, the quality model determines which quality characteristics will be considered when evaluating the properties of a software product. The product quality model categorises quality properties into eight quality characteristics illustrated in **Figure 5**.

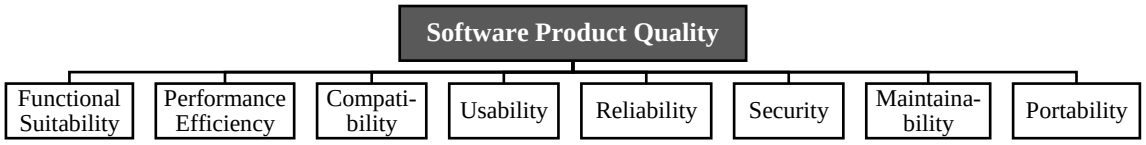

Figure 5: Software Product Quality model by ISO/IEC 25010 [46]

Each of the previous characteristics is composed of a set of related sub characteristics. The focus of the present work relies on *usability*; by that, the main interest relies on the decomposition of its sub characteristics, understood in **Figure 6**.

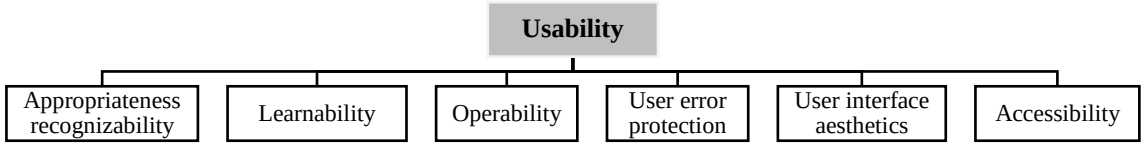

Figure 6: Usability decomposition into sub characteristics by ISO/IEC 25010 [46]

The definition of usability in sub characteristics include words that themselves require definition; detailed according to ISO/IEC 25010 [46]:

Table 4: Usability sub characteristics by ISO/IEC 25010 [46]

	DEFINITIONS
Appropriateness recognizability	"Degree to which users can recognise whether a product or system is appropriate for their needs"
Learnability	"Degree to which a product or system can be used by specified users to achieve specified goals of learning to use the product or system with effectiveness, efficiency, freedom from risk and satisfaction in a specified context of use"
Operability	"Degree to which a product or system has attributes that make it easy to operate and control"
User error protection	"Degree to which a system protects users against making errors"
User interface aesthetics	"Degree to which a user interface enables pleasing and satisfying interaction for the user"
Accessibility	"Degree to which a product or system can be used by people with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use"

1.4.1. User perceived quality and quality of use

Discussing quality compels the understanding of *user perceived quality* as the user judgement of product quality and the *quality of use*.

Quality of use is, by definition from [4], "*the extent to which a product satisfies stated and implied needs when used under stated conditions. It is determined by the particular users, tasks and environments*". Bevan defined the three measures of quality of use as *effectiveness, efficiency* and *satisfaction* – which will intersect, later, with the dimensions of usability. This relationship between quality of use and usability is determined by the measures of quality of use to be the provider of the criteria to determine whether the design of the attributes is successful in achieving usability.

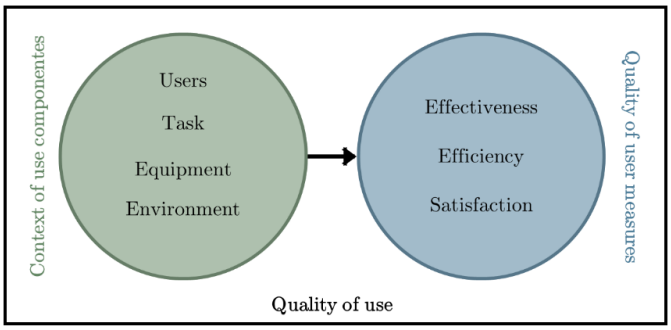

Figure 7: Usability Factors. Adapted from [5]

Thus far, quality and usability are intimately related. This is not only because usability is a software quality attribute but also because it is the quality of use in a context [44].

Table 5: Software Quality Attributes [5]

ATTRIBUTE	DESCRIPTION
Correctness	"Extent to which a program satisfies its specifications and fulfils the user's mission objectives"
Reliability	"Extent to which a program can be expected to perform its intended function with required precision"
Efficiency	"Amount of computing resources and code required by a program to perform a function"
Integrity	"Extent to which access to software or data by unauthorised persons can be controlled"
Usability	"Effort required to learn how to operate, prepare input, and interpret output of a program"
Maintainability	"Effort required to locate and fix an error in an operational program"
Testability	"Effort required to test a program to ensure that it performs its intended function"
Flexibility	"Effort required to modify an operational program"
Portability	"Effort required to transfer a program from one hardware configuration and/or software system environment to another"
Reusability	"Extent to which a program can be used in other application – related to the packaging and scope of the functions that programs perform"
Interoperability	"Effort required to couple one system with another"

1.5. Usability

Usability is a necessary condition for survival as it assesses how easy user interfaces are to use [47] in every existing interaction between humans and software or systems. Software usability is a fundamental determinant of productivity and of the acceptance of software applications [48]. The term *usability* has been defined differently through time in distinct standards gathered in **Table 6**.

Table 6: Usability definitions [49].

DEFINITIONS OF USABILITY	
IEEE Std.610.12 (1990)	"The ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component."
ISO 9241-11 (1998)	"The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."
ISO/IEC 9126-1 (2000)	"The capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions."
ISO/IEC 25012 (2006)	"The capability of the data to be understood, managed and used to catch the users attention, when used under specified conditions."

From the four standards, the fundamental assumption of the definition of usability is based on ISO 9241-11 because it was instead adopted by Human-Computer Interaction (HCI) experts and describes how to identify information for evaluating usability [49]. Considering the definitions above, usability depends on interaction to be measured and recognised on what is missing for *good usability* to be achieved.

ISO 9241-11 references three crucial issues of user interfaces, i.e. the most essential characteristics of interaction, gathered in the following **Figure 8**.

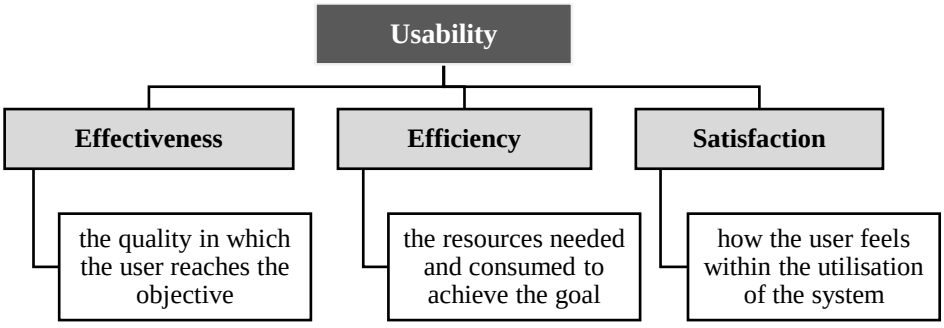

Figure 8: Usability components by ISO/IEC 9241-11[9]

Usability is not a unique property of an interface. It is intended to base usability engineering on several components that can be measured, evaluated, and tested [41]. According to author [43], usability is defined by five components detailed next, in **Figure 9**.

[9] https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en

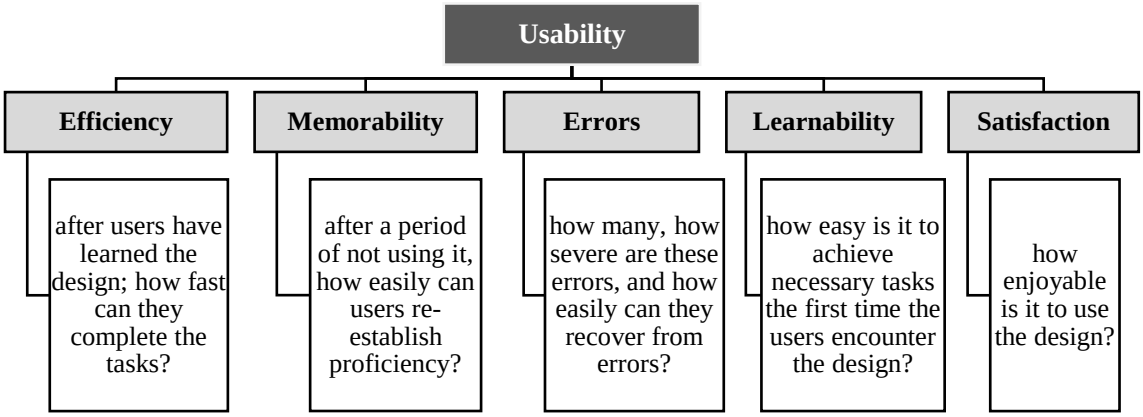

Figure 9: Nielson's usability components [43]

By now, the importance of usability is quite logical, not only because of its efforts to cover the questions mentioned above but also from the benefits it might offer: customers satisfaction, a better quality of the product, increased productivity, possible lower costs on development, user support and training [41]. Different authors might describe these components as rules or principles, but despite the label assigned, the final goals for usability are the same.

Interaction Design Foundation[10] declares that a functional and usable interface has three primary outcomes; thus, for the users, it should be easy to:

1. become familiar with the interface and be able to use it during the first contact
2. achieve their objective by using the website
3. remember the user interface and how to use it on following visits

"Usability is what determines whether a design's existing attributes make it stand or fall" [10]. Yet, the attributes that a product requires for usability are subjected to the user's nature, the task and the environment [48], i.e., the *context of use*.

1.5.1. Levels of usability metrics

The standard that concerns with software engineering and product quality, ISO/IEC 25010, refers to three different levels of quality metrics that can be transferred into the context of usability metrics [50]:

- ***Internal:*** which measure a set of static attributes and is related to software architecture. Used for predicting the extent to which the software can be understood, learned, operated and attractive.
- ***External:*** which relate to the behaviour of a system, i.e., rely on the execution of the software tested by users (at least eight users are necessary) without any hints or external assistance.

[10] https://www.interaction-design.org/

- ***In-use:*** which involve actual users in a given context of use and commonly assessed with user studies in a more controlled setting.

By this, in order to evaluate usability, the first step is to state which of the three metric levels is under investigation.

1.5.2. Usability measures

The importance of measuring usability relies on the high utility since it quantifies how well users can interact with a certain product or service [51] as well as to obtain a more complete understanding of users' needs to improve a product and consequently provide a better user experience [52]. Measuring the usability of a product or service defines the accomplishment or not of its purpose. Even if a product performs its primary technical function perfectly but the user cannot get the product to work, then that product has failed [51].

According to ISO/IEC 25023 the aim of usability measures is *"to be able to measure the degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"*. The standard defines the measures of usability as the sub characteristics previously detailed in **Figure 6**.

Quoting Nielsen, *usability plays a role in each stage of the design process*, and it should be considered since the beginning until the end of the new design. It is challenging to specify precisely the measurable usability attributes and their interpretations from different perspectives because it measures three dimensions at the same time: user performance, satisfaction, and productivity [53].

Inherent to the first three previous usability features (effectiveness, efficiency and satisfaction), ISO 9241-11 suggests that the measurement of usability can be based recurring to the following questions, respectively [48]:

- How successfully do users achieve their goals using the system?
- What resources are spent to achieve their goals?
- How do the users feel about using the system?

The measures of effectiveness are explained by [41] as to be related with the goals of the user regarding the quality and completeness with which these objectives can be achieved, efficiency measures relate the effect achieved with the expenditure of resources and satisfaction measures the way the users feel while using the system.

1.5.3. Evaluation

Evaluation, by definition from [54], is a "*systematic determination of the extent to which an entity meets its specified criteria*".

When specifying or evaluating usability, it is important that the context selected adequately represents the important aspects of the actual or intended context of use [52]. Some authors in turn, state that, in the phase of evaluation, not only usability needs to be evaluated, but also functionality and the

users' experience and additionally, identify specific problem of the design (in terms of functionality and usability) in order to be corrected on the next development cycle [41]. Complementing these affirmations, evaluation involves, and that goes from observing users, asking users, asking experts, user testing, inspections to modelling users' performance [55].

1.5.3.1. Types of evaluation

Now that the concept of evaluation and its main objectives is clearer, it is time to go through the different types of evaluation.

Table 7: Types of evaluation

Analytical Evaluation (without users)	**Inspection Methods**	Expert plays the user's role, tries to identify usability issues
	Predictive Methods	Analyses physical and mental operations to achieve a specific task
Empirical Evaluation (with users)	Measure the performance and satisfaction of users performing tasks with performance indicators	
Formative Evaluation	Performed during the iterative cycle of the design	
Summative Evaluation	Performed at the end of the final design	

When evaluating a system in terms of usability, it makes sense to understand several characteristics of the future users of the product itself. In this sense, there are three dimensions truly important to be considered when evaluating a system or product [41]:

- Functionality
- Usability
- User experience and the impact it has on users

In order to deliver a system abiding by all the previous points, many studies have been made with different methods and techniques to evaluate systems and products. The following subchapter lists those methods gathered during this investigation.

1.5.3.2. Evaluation methods and techniques

Diving in ISO 25060 [56], three evaluation techniques were found:

- **Inspection:** inspectors review the interface and evaluate each element of the interface against a list of commonly accepted principles or heuristics or specified user requirements. Inspection techniques can be further characterised as heuristic evaluations, formal usability inspections, feature inspections, consistency inspections, standards inspections, and guideline checklists;
- **Walkthroughs:** evaluators construct task scenarios from a specification or context of use description or early prototype and after role-play the part of a user working with that interface. Walkthroughs include cognitive walkthroughs and pluralistic walkthroughs (a team consisting of an expert evaluator working with users (end users, developers...) as a group);

- **Usability Testing:** is about observing a sample of users performing specific given tasks with a system. A major component of usability testing is that the user is expected to complete tasks with minimum assistance in order to enable measurements of effectiveness, efficiency, and user satisfaction.

The previous concepts helped the understanding of previous literature on the matter of evaluation methods and techniques. From the studies' review, seven different methods were found to be capable of performing a reliable usability test on systems or products. Those methods are summarised in **Table 8** with the correspondent authors of the papers. From that analysis, questionnaires are the most popular method among the different researchers, followed by recording methods.

Table 8: Usability evaluation methods

Author(s)	Participants			Methods						
	#	Experts	Users	Heuristic	Recording	Observation	Focus Group	Questionnaires	Think-Aloud	Interviews
[57]	21		•					✓		
[55]	58		•					✓		
	4	•		✓						
[58]	115		•					✓		
[32]	6		•		✓					✓
	22							✓		
[59]	7		•		✓			✓	✓	
[60]	3		•		✓					✓
[42]	10	•					✓			
[61]	151		•					✓		
[62]	105		•					✓		
	7		•		✓					
[63]	12	•						✓	✓	✓
[64]	80		•			✓	✓	✓		
[6]	98		•					✓	✓	
[65]	22		•			✓	✓	✓		

The different methods and techniques found in the literature review consist of the interpretations gathered in the following paragraphs.

Heuristic

The heuristic evaluation is a way to evaluate a user interface in order to find usability problems in the design of the user interface to correct them forwardly [52]. It is usually performed by expert evaluators; therefore, it is useful when the study focuses on expert feedback rather than users' feedback.

Recording

This evaluation method applies to record audio and/or video. Recording the participants enable the collection of essential data, such as participants' reactions (e.g. facial expressions) and interactions (e.g. to a product/software in the usability test.

Recording can be facilitated through automated usability testing tools [66]. An example of one of these automated tools is the *Morae* program [67] that allow researchers to replay the recorded data and mark the users' reactions on a timeline [59].

Observation

User observation method can be qualitative, observing the behaviour of the users to identify actual usability problems, or quantitative, measuring user performance and responses to obtain data on two usability components: effectiveness and efficiency [68]. An advantage of using this method is to assess if the participants take the expected actions or movements towards the system/product in testing.

Focus group

Focus group is identical to group interviews but the fundamental difference relies on the role adopted by the researcher and the relationship with the participants, i.e., while in interviews the researcher plays the role of an *investigator* (centre-stage), in a focus group discussion, the researcher adopts the role of a *facilitator* or *moderator* (peripheral) [69].

Questionnaires

Although this method takes place in most of the papers reviewed, as can be seen in **Table 8**, it is always accompanied by another method. Questionnaires are acknowledged for collecting demographic data and user's opinions [55]. Over time, several questionnaires have been developed to evaluate usability dimensions. The most well-known questionnaires, also referred as Post-Study Questionnaires because are administrated after users test the products, for studying usability are the following:

- Computer System Usability Questionnaire (CSUQ)
- Post-Study System Usability Questionnaire (PSSUQ)
- Questionnaire for User Interaction Satisfaction (QUIS)
- Software Usability Measurement (SUMI)
- Software Usability System (SUS)
- Usability Metric for User Experience (UMUX and UMUX-Lite)
- Usefulness, Satisfaction, and Ease of use (USE)

The following **Table 9** summarizes a comparison between the questionnaires previously listed.

Table 9: Comparison of Post-Study Questionnaires (based on a broad adaptation of [21])

Scale name	# of items	# of subscales	Scale type	Reliability (Chronbach's alpha)	Evidence for validity	Evidence for sensitivity	Studies
CSUQ	16	3	Likert (7) + N/A option	.94	Yes	Yes	[51]
PSSUQ	16	3	Likert (7) + N/A option	.94	Yes	Yes	[70]
QUIS	27	5	Bipolar (9)	.94	Yes	Yes	[71]
SUMI	50	5	Likert (3)	.92	Yes	Yes	[72]
SUS	10	2	Likert (5)	.89	Yes	Yes	[51]
UMUX	4	3	Likert (7)	.91	Yes	Yes	[51]
UMUX-Lite	2	-	Likert (7)	.94	Yes	-	[73]
USE	30	4	Likert (7) + N/A option	.98	-	-	[61]

Think aloud

Thinking aloud or concurrent verbalisation was borrowed from cognitive psychology [74]. This method consists of users to think aloud while performing a set of specified tasks [7], in other words, to verbalise anything that crosses their mind during the task performance [74]. An advantage of this method is that it enables the collection of insights on the difficulties that participants encountered while using the system/product [59].

Interviews

Interviews enable the collection of subjective information from users or experts. Some authors [68] sort this information into qualitative (problems, opinions or impressions given during or after the usability evaluation) or quantitative (measures of user satisfaction or perception). An example of what to ask participants in an interview can be why did they take a particular action or made a particular choice. And a type of interview can be a stimulated recall interview, which intention is to help participants recall specific moments during the test [32].

1.5.4. User experience

The concept of *user experience (UX)* is broadly used but understood differently [75]. There are many definitions for user experience that were gathered from the literature and the web[11], but taking ISO 9241-210 definitions of user experience: "*a person's perceptions and responses that result from the use or anticipated use of a product, system or service*" [12].

When speaking of user experience, a main confusion of what is the role of usability linked with user experience ascends. These two concepts are overlapping concepts and intimately related not only

[11] http://www.allaboutux.org/ux-definitions
[12] https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-1:v1:en

because UX work implies usability measurement but also because of the relationship between usability goals and user experience goals, shown in **Figure 10**.

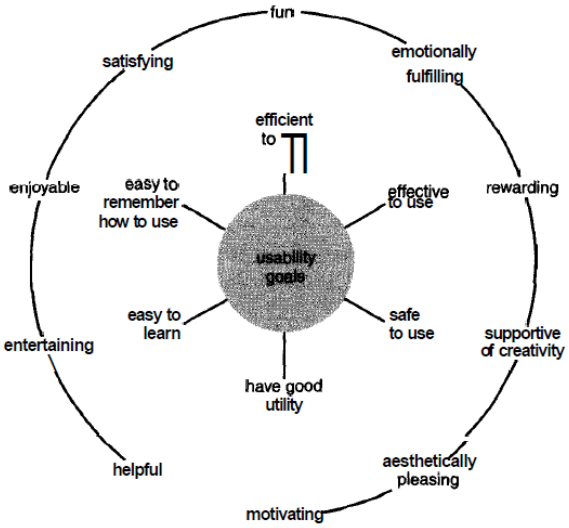

Figure 10: Usability and user experience goals [76]

Usability goals are operationalised through specific criteria and user experience goals are less clearly defined but concern with explaining the quality of the user experience [76].

1.6. Related work

Although it is a recent area of research, there is related work worth mentioning that will aid to enrich and frame the developments described in the following chapters. Following are summaries of studies on software products and dashboard usability.

1.6.1. Usability testing

Usability testing discusses to the evaluation of a product or system by testing it with users, preferably the end users of the product or system in evaluation. The goal is to identify usability problems and collect useful data and the user's satisfaction on the product. Several studies have been read and analysed on how authors along the years have been developing usability tests, planning tests, and analyse the findings. Following, summaries of the most relevant work to follow this investigation are presented.

A study of testing usability of a mobile app prototype – *MyLA* app and a web browser version as well was conducted through a mixed-method design and counted on the participation of 105 students, of which 7 joined an additional eye tracking study with three tasks to solve [62]. The authors divided the usability test into two parts:

1. Getting the participants familiar with the app prototype and
2. Respond an online questionnaire

The MyLA app has three main categories: *My Profile, My Learning, My Progress.* It also has a dashboard for students with the home page containing: *Pinboard Entries* (managing messages), *Survey*

Center (managing surveys) and the *LectureTracker* (for instructors to observe the aggregated values of the students). *MyLA* usability testing resulted in the following statements: (1) the handling of the *MyLA* Web app prototype is intuitive; (2) the app's structure is easy to learn; (3) the navigation within the app is clear and user-friendly (4) the students mostly like the idea of *MyLA*; (5) the design and colours can be improved, because the opinions deviate fairly high.

An evaluation usability model for measuring the usability of mobile office applications user interfaces was found in the literature [77] using a model that is intended for smartphones. The authors measured the effectiveness, usefulness and reliability through two office applications. The model was developed within three layers: usability factors, usability sub factors and measurable criterion, respectively. The definition of the components of each layer were based in usability assessment of previous literature review.

The authors used two case studies for usability testing conducted with 65 participants. The model suggested was validated by user centred methods: usability testing, think-aloud technique during pilot testing and questionnaires.

The conclusions to outline are that the model enables application designers to guarantee a more adept and enhanced usability of office applications, contributions with a collection of metrics for smart phones office apps and the model will facilitate users by interacting with more user-friendly applications.

Regarding quantifying usability, a paper was found where the authors [78] evaluated a diabetes mHealth system on three usability dimensions: efficiency, effectiveness and satisfaction. To measure efficiency and effectiveness, the authors used metrics presented in the International Organization for Standardization 9241-11 standard. To measure satisfaction, the SUS Questionnaire was the instrument chose along with the Morae software to measure the other dimensions. The usability testing had the participation of 10 randomly chosen patients. The participants received standardised training and had to perform representative tasks and fill in a pre-test survey. From the data provided by the participants, the authors gathered the following results: (1) the average SUS score indicated good but not excellent system usability; (2) male par were more successfully in task completion; (3) younger participants had higher performance scores; (4) patients with more experience in information technologies had higher performance rates.

1.6.2. Dashboards' usability testing

Along with the literature research, several studies about measuring dashboards' usability were found. Studies conducted in health, energy and education. Next, summaries of the most relevant and related work are presented. The following studies/experiments were chosen considering the criterion of the methodology used and similarity to this investigation.

A relevant paper, from [79], presented the designing and development of a framework to measure a dashboards' usability and efficiency, adapting the ISO/IEC 9241-11 and based on the principles of BI and HCI design. The project was held with professionals in both areas from a Portuguese company, HMI teachers and students with knowledge in both areas too. The dashboard was applied in a case study in the energy sector.

Dashboards do bring several advantages; however, the authors perceived that these results are not always achieved due to lack of usability and efficiency of some interfaces. The main goals of their experiment were to verify, through hypotheses, usability differences on different versions of the dashboard according to the number of errors introduced. Conclusions outlined: (1) the usability of the error-free dashboard is higher compared to dashboards with five to seven errors; (2) the usability differs significantly from a solution without errors than solutions with five to seven errors; (3) compared to professionals, students attributed lower usability.

Authors [63] conducted a usability study that requires lesser time than usual but still useful and efficient. It aimed to firstly evaluate the usability of an interactive surgical dashboard and find quality improvement opportunities, and secondly demonstrate the mismatched mental models between the users of the dashboard and its designers. Their two-phased study was based on the comparison of an interactive dashboard that provides real-time and interactive access to data from clinical outcomes for the surgical program and an existing static, spreadsheet-based dashboard - conducted with interviews with the designers as the first phase and questionnaires, about both existing new dashboards, to the users as the second phase.

The new dashboard got a SUS score of 82.9 against the 63.5 on the existing dashboard. They also highlight the importance of conducting a user-centred evaluation that identifies the gaps between designers' intentions and users' mental models because even though designers' features did provide a solution for the problem, at the same time introduced several usability issues and identified areas of improvements from their testing.

Another relevant paper is a study on the usability evaluation of a dashboard designed for home care nurses [80]. The dashboard was developed under a *design science framework* to support decision making in the care of patients with heart failure.

They conducted two initial studies with the nurses: to find what information was required to assist their patients and to explore nurses' ability to understand visualised information. The design process included nurses' feedback in all aspects. Evaluation methods used: Tasks, Users, Representations, and Functions (TURF) framework by [81], analysing participants data, heuristic evaluation and task analysis (time on task comparison). The questionnaires used the 10-item System Usability Scale and 50-item Questionnaire for User Interaction Satisfaction. The dashboard presented minor usability issues within flexibility and efficiency.

Chapter 2:

Model Proposal

"If you cannot measure it, you cannot improve it."

(Lord Kelvin, a.k.a. Sir William Thomson, n.d.)

This chapter provides an overview of the development process of a new model to evaluate usability in dashboards, the Dashboard Assessment Usability Model (DATUS).

In interactive software systems, usability is increasingly recognized as an important quality factor [6][3]. Measuring usability is still limited to companies since they still under-invest in usability due to expensive metrics, ways to measure them and because usability testing is a lengthy process. Despite that, quality is determinant to disclose the success of the product or software developed among end-users. Although there are many methods for evaluating usability in products and software systems, when it comes to dashboards, there is no clear method or model to assess its usability. As stated before, this gap led to the research carried out in this dissertation, that is, a proposal of a new model, accompanied by an evaluation method, that can be applied to the usability assessment of dashboards.

The following pages describe the development of the proposal taking into account the three main goals of usability evaluation [55]:

- measurement of the extent and accessibility of the dashboard functionality
- measurement of the user's experience of the interaction
- identification of any particular issues within the dashboard

The process of the development was twofold:

1. Instrument identification: a review of previous studies and related work to generate an initial set of dimensions and correspondent measurement items, i.e. metrics;
2. Instrument confirmation: definition through discussion of the dimensions to take part in the model proposal with the correspondent definitions and metrics.

2.1. Evaluation dimensions

It is difficult to define the nature of the features and attributes required for the evaluation of a product's usability since it depends on the context in which the product is used [5]. The development of the DATUS model considered the broadest applications where a dashboard can be applied, that is, no specific context of use was considered; hence, the proposed model can be applied to any dashboard. DATUS is hierarchical, decomposing usability into dimensions, then into detailed metrics.

From the analysis of the thirty-one dimensions found in the literature, DATUS comprises eight usability dimensions for evaluation: effectiveness, efficiency, satisfaction, learnability, accessibility, appropriate recognizability, user interface aesthetics, and operability. The acceptance criterion for these

dimensions started from analysing the definitions of each from the International Organization for Standardization plus further descriptions given from the different authors and finished with additional discussion. Following, each dimension was given an adapted definition taking into account the scope of dashboards. The detail of each usability dimension in the DATUS model is described in **Table 10**. It is also worth mentioning the reasons why certain dimensions, out of the 37 found in the literature, were discarded and **Annex A** summarises the discussion process that took place for each of the 37 dimensions, and lists the reasons for exclusion.

Table 10: DATUS usability dimensions

Dimension	Definition	Source	DATUS Adaptation
Effectiveness	*Accuracy and completeness with which users achieve specified goals*	ISO 9241-11:2018 - Usability: Definitions and concepts	Determines if a user can achieve specified analytical goals with precision and perfection, therefore removing information from the dashboard with these attributes; if it is fit for purpose (real usefulness)
Efficiency	*Resources used in relation to the results achieved*		Determines if the dashboard fulfils the planned objectives considering the resources spent (time, effort and computational resources)
Satisfaction	*Extent to which the user's physical, cognitive and emotional responses that result from the use of a system, product or service meet the user's needs and expectations*		Determines how satisfied are the user with the dashboard, if it meets their needs, that is, the overall analytical goals
Learnability	*Degree to which user needs are satisfied when a product or system is used in a specified context of use*	ISO 25010:2011 - Systems and software Quality Requirements and Evaluation	Determines how easy it is to read and understand, it comprises the degree in which a user can learn to use the dashboard quickly and achieve the expected analytical goals without too much help
Accessibility	*Degree to which a product or system can be used by people with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use*		Determines if the dashboard is accessible to all users, even with some kind of disability, that is inclusive
Appropriate recognizability	*Degree to which users can recognize whether a product or system is appropriate for their needs*		Determines if users recognize whether the dashboard is appropriate for their needs
User interface aesthetics	*Degree to which a user interface enables pleasing and satisfying interaction for the user*		Determines the visual attractiveness of the dashboard, in other words, if it is visually attractive to interest a user in interacting with the dashboard
Operability	*Degree to which a product or system has attributes that make it easy to operate and control*		Determines if the dashboard follows the Business Intelligence and visualisation practices, allowing the removal of relevant information for decision making

2.2. Evaluation metrics

Along with the first list of dimensions found in several papers, there was also a collection of possible metrics already used by other authors to measure the dimensions chosen for the model (see **Annex E**). The analysis enabled reaching the final association of the metrics with the respective dimensions of the model, as represented in the following **Table 11** and the complete model DATUS in **Figure 11**.

Table 11: DATUS metrics

Dimension: Effectiveness					
#	**Metric**	**Unit**	**Description**	**Formula**	
1	Number of analytical goals completed successfully	Number	Number of tasks finished with the correct outcome	$\frac{\Sigma\, number\ of\ tasks\ with\ the\ correct\ outcome}{\Sigma\, number\ of\ tasks}$	(2.1)
2	Number of errors	Number	Average number of incorrect outcomes given by users when trying to achieve a goal	$\frac{\Sigma\, number\ of\ tasks\ with\ the\ wrong\ outcome}{\Sigma\, number\ of\ tasks}$	(2.2)
3	Perception of analytical goals achievement	Likert Scale	Degree to which the dashboard allows users to complete all desired analytical goals	-	
4	Data trustfulness	Likert Scale	Indirect metric regarding trustfulness: assessment of the current date and date of data extraction is presented in the dashboard	-	
Dimension: Efficiency					
#	**Metric**	**Unit**	**Description**	**Formula**	
5	Time to complete a goal	Seconds	Total time, in seconds, to execute a given goal	-	
6	Perception of time to complete a goal	Likert Scale	Degree of which the user is able to complete his work using the dashboard	-	
7	Perception of navigation ability	Likert Scale	Degree of which the user is able to efficiently complete his work using the dashboard	-	
8	Number of actions to complete a task	Number	Number of clicks to complete a task	$\Sigma\, clicks\ per\ task$	(2.3)
Dimension: Satisfaction					
#	**Metric**	**Unit**	**Description**	**Formula**	
9	Overall satisfaction	Likert Scale	Overall user perception of how useful, usable and likeable the dashboard is	-	
10	User experience while performing given tasks	Likert Scale	User individual perception and thoughts about his performance concerning the dashboard	-	

Dimension: Learnability				
#	**Metric**	**Unit**	**Description**	**Formula**
11	Learning curve	Likert Scale	Degree to which the dashboard is easy/simple to learn how to use and the user becomes productive quickly	-
12	Task time of the first use	Seconds	Total time, in seconds, for users to execute a given list of tasks for the first time	-
13	Task time of the second use	Seconds	Total time, in seconds, for users to execute a given list of tasks for the second time	-
Dimension: Accessibility				
#	**Metric**	**Unit**	**Description**	**Formula**
14	Readability	Likert Scale	Degree to which all features of the dashboard are readable to a wide range of users	-
Dimension: Appropriate recognizability				
#	**Metric**	**Unit**	**Description**	**Formula**
15	Fitness for purpose or use	Likert Scale	Degree to which users can recognise whether the dashboard is suitable for their needs	-
Dimension: User interface aesthetics				
#	**Metric**	**Unit**	**Description**	**Formula**
16	Visual attractiveness	Likert Scale	Degree to which the dashboard is visually appealing for the user	-
Dimension: Operability				
#	**Metric**	**Unit**	**Description**	**Formula**
17	Visual hierarchy of data	Likert Scale	Degree of which the user can easily understand the needed and most critical information	-
18	Consistency	Likert Scale	Degree of which the dashboards enables consistent understanding by different people	-
19	Simplicity	Likert Scale	Degree of which the dashboard presents complex information in an understandable and simpler form	-
20	Recoverability	Likert Scale	Degree of each the user has the ability to go back to a previous state of the analysis with the dashboard	-

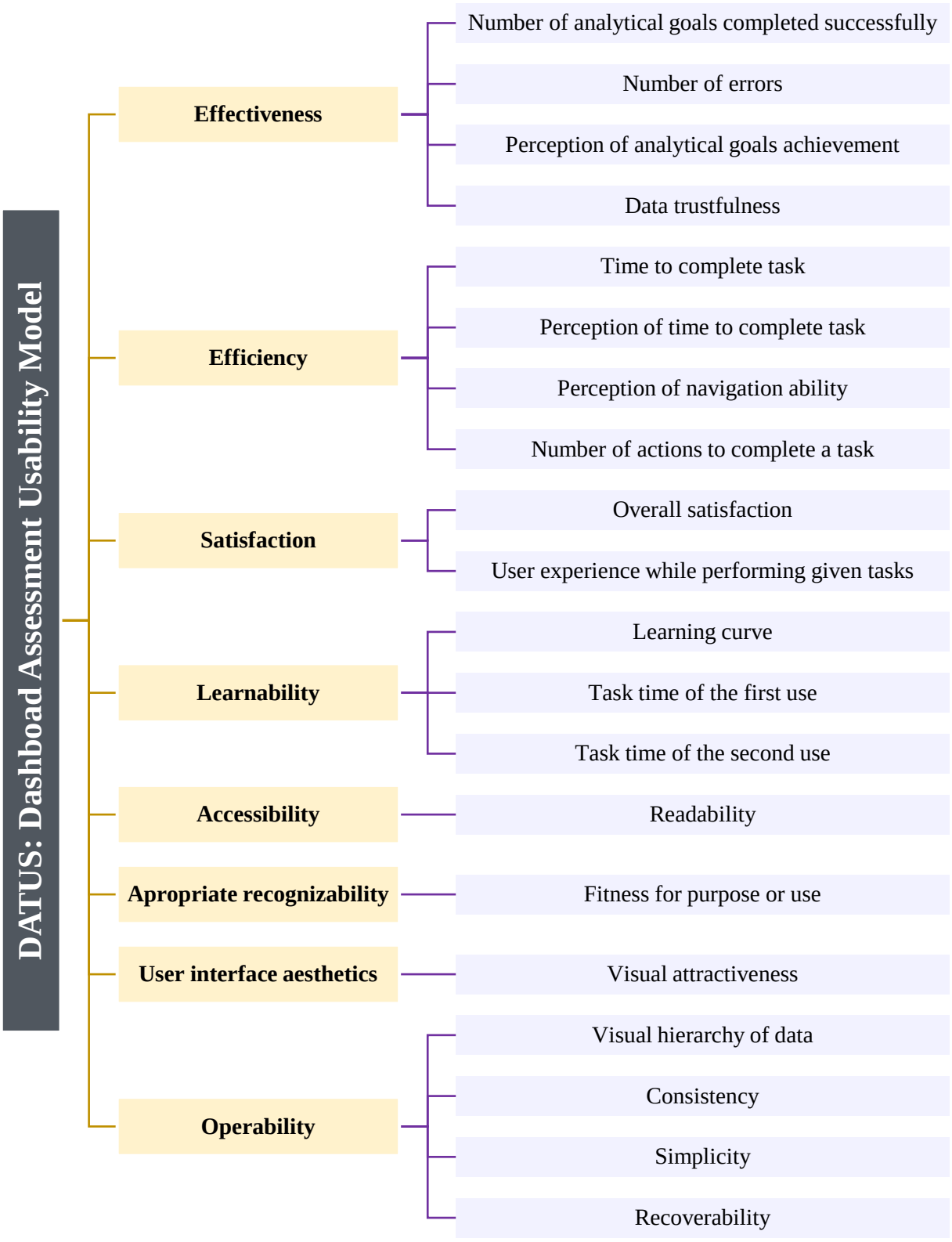

Figure 11: DATUS model proposal hierarchically organized in dimensions and metrics

2.3. Evaluation methods

DATUS foresees the evaluation of 20 metrics distributed over eight dimensions. Two methods were selected to assess the dimensions: recording and questionnaire; both were identified in Section 1.5.3.2, as the most popular among the review of different studies. The correspondence between the unit in which the metric must be measured and the method to be applied is shown in **Table 12**.

Table 12: Methods to measure DATUS metrics

Dimension	#	Metric	Unit	Method
Effectiveness	**1**	Number of analytical goals completed successfully	Number	Recording
	2	Number of errors	Number	Recording
	3	Perception of analytical goals achievement	Likert Scale	Questionnaire
	4	Data trustfulness	Categorical	Recording
Efficiency	**5**	Time to complete a task	Seconds	Recording
	6	Perception of time to complete a task	Likert Scale	Questionnaire
	7	Perception of navigation ability	Likert Scale	Questionnaire
	8	Number of actions to complete a task	Number	Recording
Satisfaction	**9**	Overall satisfaction	Likert Scale	Questionnaire
	10	User experience while performing given tasks	Likert Scale	Questionnaire
Learnability	**11**	Learning curve	Likert Scale	Questionnaire
	12	Task time of the first use	Seconds	Recording
	13	Task time of the second use	Seconds	Recording
Accessibility	**14**	Readability	Likert Scale	Questionnaire
Appropriate recognizability	**15**	Fitness for purpose or use	Likert Scale	Questionnaire
User interface aesthetics	**16**	Visual attractiveness	Likert Scale	Questionnaire
Operability	**17**	Visual hierarchy of data	Likert Scale	Questionnaire
	18	Consistency	Likert Scale	Questionnaire
	19	Simplicity	Likert Scale	Questionnaire
	20	Recoverability	Likert Scale	Questionnaire

2.3.1. Recording and analysis

The recording method is a passive usability testing method to provide insights into how users interact with the dashboard in their "natural environment" in order to spot major problems with the dashboard projected functionality. This method is also accompanied by a further analysis of the data that is collected from the user assessment tasks. In DATUS it must consist of recording the computer screen of the participants and analysing the following:

- Number of analytical goals completed successfully;
- Number of errors;
- Number of actions to complete a task;
- Task time of the first use;
- Task time of the second use;
- Current date and date of data extraction/Current or date of data extraction/None

The last item listed previously intends to evaluate the indirect metric regarding trustfulness of the data that is obtainable, that is, assess if the current date and date of data extraction are presented in the dashboard.

2.3.2. Questionnaires

Questionnaires are *a measurement tool designed to assess a computer user's subjective satisfaction with the human-computer interface* [70].

Six questionnaires were analysed: CSUQ, QUIS, SUS, UMUX, SUMI and USE in order to find the questionnaire that would be used with the DATUS model. For each questionnaire, a mapping was made between the dimensions of the proposed model and the subscales or groups from the existing questionnaires (for instance, usable or learnable from SUS). None of the six analysed questionnaires measures all the metrics of the DATUS model. Therefore, there was a need to design a DATUS questionnaire (**Table 13**) to support the evaluation of the model. The goal was to build the DATUS questionnaire with as much as possible existing questionnaires, given the existing evidence of construct validity and levels of reliability already performed in previous studies.

The adaptation consisted on the merge of 18 questions (1 to 18 in DATUS questionnaire, **Table 13**) from The Computer System Usability Questionnaire (CSUQ) and two questions (19 and 20 in DATUS questionnaire, **Table 13**) from the Questionnaire for User Interface Satisfaction (QUIS). The full versions of the original questionnaires are available in **Annex B**. The merge between two existing and well-studied questionnaires, implied modifications in the CSUQ and QUIS questions such as:

- Scale changes: from a 7-point to 5-point Likert Scale and reversed direction of the anchors. These adaptations were sought to fasten the evaluation;
- Adaptation of the original sentences to a dashboard context;
- Definition in subscales adapted to definition in dimensions of the model;

The final DATUS questionnaire, presented in **Table 13**, is a 20-item survey aligned with a five-point Likert Scale (1 = "Strongly disagree"; 5 = "Strongly agree" plus Not Applicable = "NA" option). Additionally, two open-ended questions are also available for participants to express comments and suggestions:

1. Is there any information you would like to see that has not been considered on the dashboard?
2. Do you have any other comments or suggestions you want to share with us?

Table 13: Questionnaire for DATUS

#	Questions	Strongly disagree				Strongly agree	
		1	2	3	4	5	NA
1	Overall, I am satisfied with how easy it is to use this dashboard.	○	○	○	○	○	○
2	It was simple to use this dashboard.	○	○	○	○	○	○
3	I can effectively complete my work using this dashboard.	○	○	○	○	○	○
4	I am able to complete my goals (tasks) quickly using this dashboard.	○	○	○	○	○	○
5	I am able to efficiently complete my goals (tasks) using this dashboard.	○	○	○	○	○	○
6	I feel comfortable using this dashboard.	○	○	○	○	○	○
7	It was easy to learn to use this dashboard.	○	○	○	○	○	○
8	I believe I became productive quickly using this dashboard.	○	○	○	○	○	○
9	Whenever I make a mistake using the dashboard, I recover easily and quickly.	○	○	○	○	○	○
10	The information (on-screen messages) provided with this dashboard is clear.	○	○	○	○	○	○
11	It was easy to find the information I needed.	○	○	○	○	○	○
12	The information displayed in the dashboard is easy to understand.	○	○	○	○	○	○
13	The information displayed in the dashboard is effective in helping me complete the tasks and scenarios.	○	○	○	○	○	○
14	The organization of the information on the dashboard is clear.	○	○	○	○	○	○
15	The interface of this dashboard is pleasant.	○	○	○	○	○	○
16	I like using the interface of this dashboard.	○	○	○	○	○	○
17	This dashboard has all the functions and capabilities I expect it to have	○	○	○	○	○	○
18	Overall, I am satisfied with this dashboard.	○	○	○	○	○	○
19	Data on the dashboard is easy to read.	○	○	○	○	○	○
20	Visual encoding of data is consistent throughout the dashboard.	○	○	○	○	○	○

Assessing usability through questionnaires is a common method among usability studies. As described before, well- known questionnaires have been analysed to understand if there would be one which could measure all the metrics defined for DATUS.

To assess the usability of the prototype using the DATUS questionnaire, the CSUQ rules for calculating the scores were considered and adapted [70]:

1. Average of each question and the average of each dimension of the DATUS model. This rule is derived from the CSUQ rule which says *average the scores from the appropriate items to obtain the scale of each questions and subscale scores*, and instead of subscales we have dimensions;
2. High scores are better than low scores due to the anchors (Strongly disagree (1) and Strongly agree (5)) used in the 5-point scales;
3. If a participant marks "N/A" or does not answer an item, then average the remaining item scores.

Table 14 shows the correspondence between each dimension that in being evaluated with the question from the DATUS questionnaire.

Table 14: Mapping between DATUS questionnaire and dimensions.

Dimension/Group	#	Question
Satisfaction	**1**	Overall, I am satisfied with how easy it is to use this system.
	6	I feel comfortable using this system.
	16	I like using the interface of this system.
	18	Overall, I am satisfied with this system.
Effectiveness	**3**	I can effectively complete my work using this system.
	13	The information is effective in helping me complete the tasks and scenarios.
Efficiency	**4**	I am able to complete my work quickly using this system.
	5	I am able to efficiently complete my work using this system.
Operability	**2**	It was simple to use this system.
	9	Whenever I make a mistake using the system, I recover easily and quickly.
	11	It is easy to find the information I needed.
	14	The organization of information on the system screens is clear.
	20	Visual encoding of data is consistent throughout the dashboard.
Learnability	**7**	It was easy to learn to use this system.
	8	I believe I became productive quickly using this system.
	10	The information (such as online help, on-screen messages, and other documentation) provided with this system is clear.
	12	The information provided for the system is easy to understand.
User interface aesthetics	**15**	The interface of this system is pleasant.
Appropriate recognizability	**17**	This system has all the functions and capabilities I expect it to have.
Accessibility	**19**	Characters on the dashboard are easy to read.

Chapter 3:

Student dashboard design

"A dashboard needs to be light, focus on «K» in KPIs."

(Pearl Zhu)

This chapter provides an overview of the development of a student dashboard proposal for Iscte – Instituto Universitário de Lisboa.

We have been dealing with the phenomenon of information overload, which will likely prevail as long as the field of information and communication technologies and fast-paced business environments continues to progress, increasing the amount of generated and consumed data [82]. Users are provided with loads of information, and there is a need for proper tools to represent it and to avoid the negative consequences of data misinterpretation. Dashboards are one of the possible solutions, enabling the analysis of data sets through meaningful visualisations. Well-designed dashboards can improve the decision-making and support cognition and perception of users [83]. A dashboard [84] *"is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance"*.

Dashboards are not restricted to a single area of implementation. When bringing this solution to the education context, which deals with rich data sets from the e-learning and other educational platforms, we are talking about learning dashboards.

The benefits dashboards bring to students are listed depending on the data that is being treated and displayed (monitorisation, evaluation or prediction). The focus of developing the present prototype is not only to deliver a valuable artefact of learning analytics to the institution but also:

- To help student self-regulate their own learning through an analytical environment;
- To improve support for students;
- To help students make decisions and outline their college path;
- To improve institutional data;
- To foster student engagement/involvement with the courses' activities and autonomous work.

The development of the dashboard proposal started with requirements traceability from the literature review, consisting of collecting requirements and the most valuable features for student dashboards from previous studies, which can be seen in **Annex C.**

The final list consists of features from not only the literature review and defined taking into account the exclusion criteria, but also the addition of new ones considered important and not found in the previous student dashboards review.

The features raised were revised and filtered taking into account which ones could be linked with the main university platform, the academic system called Fénix. The end result is a list of 20 features from previous studies to be displayed in the prototypes, displayed in **Table 15**. The filtering was necessary due to the context in which the dashboard was designed. The goal was to produce dashboards prototypes for the academic system Fénix, that rely on the academic context for bachelor students from Iscte to monitor their performance while attending university. It is not the aim of the prototypes to support teachers or tutors to monitor their students' performance (as for [85]). Moreover, the dashboards were tailored to the needs of undergraduate students. Graduate programs, such as master and PhD programs, have other specific characteristics that have not been considered (e.g., dissertation/thesis, choice of elective courses).

Table 15: Selected features for the case study dashboards

Group	Feature-id	Feature	Data type
General Information	1	Personal information (student number and name (first and last), photography)	Qualitative
	2	Programme name	Qualitative
	3	Number of curricular units	Quantitative
	4	Year of first enrolment	Qualitative
Grades and performance	5	Length of bachelor in years (prediction with historical data)	Quantitative
	6	Current average programme grade	Quantitative
	7	Credits achieved	Quantitative
	8	Number of enrolled courses/Approved courses	Quantitative
	9	Overall student performance compared with peers	Quantitative
	10	Grades per course	Quantitative
	11	Tips as tricks from former students to help students improve their learning study behaviour	Qualitative
	12	Concluded and failed courses	Qualitative
Assessments	13	Calendar with evaluation dates	Qualitative
	14	Reminders of upcoming assignments	Qualitative
Classes	15	Attendance (indicating if the student is at risk of failing or not - only applicable for courses with a minimum number of class attendance)	Quantitative
Online data	16	Number of times the student accessed the Fénix platform compared with peers	Quantitative
	17	Number of times the student accessed the e-learning platform compared with peers	Quantitative
Others	18	Library status (requisitioned books and due dates, number of requisitioned books compared with peers)	Qualitative
	19	Tuition fees status (up to date or late)	Qualitative
	20	Unanswered/answered pedagogical surveys	Quantitative

The rationale for designing two dashboards is related to the temporal analysis of student performance, i.e., two periods of report: *concluded* versus *running*. This temporal analysis is displayed in two tabs: academic year (overview of the whole academic year completed, reporting the performance of the concluded year) and semester (the running first or second semester, showing contextual

information about what is happening throughout the semester). The features were associated with each view and new features, apart from the literature review, were added. Features are represented by widgets, i.e., graphical elements of a dashboard to represent a specific type of information. An important step of the development of the prototypes were the definition of which widgets should characterise each feature. This is important because it represents the first impression of users and plays an essential role in the usability and acceptability of the system [83]. The widgets started as drafts on paper (*sketching*) and then were translated into a digital representation using Microsoft® PowerPoint®.

The last step of the design was to split the widgets to the respective views, organised in three tabs:

- *Academic Year view* can be seen in *Ano Letivo* (Academic year in Portuguese), see **Figure 12**;
- *Semester view* can be seen in two tabs according to each semester of the current year: *1º Semestre* and *2º Semestre* (1st Semester and 2nd semester, respectively in Portuguese), see **Figure 16**.

Sequential iterations were made during the design process, originating several versions to better fit each data representation inside the correspondent view, until reaching the final proposals. The dashboard is intended to be displayed on the web page of the Fénix platform, on the Dashboard tab as illustrated in the following figures, linked to a database that allows the report of the information to be constantly updated.

All data presented in the next figures is exclusively for illustration purposes.

3.1. The Academic Year view

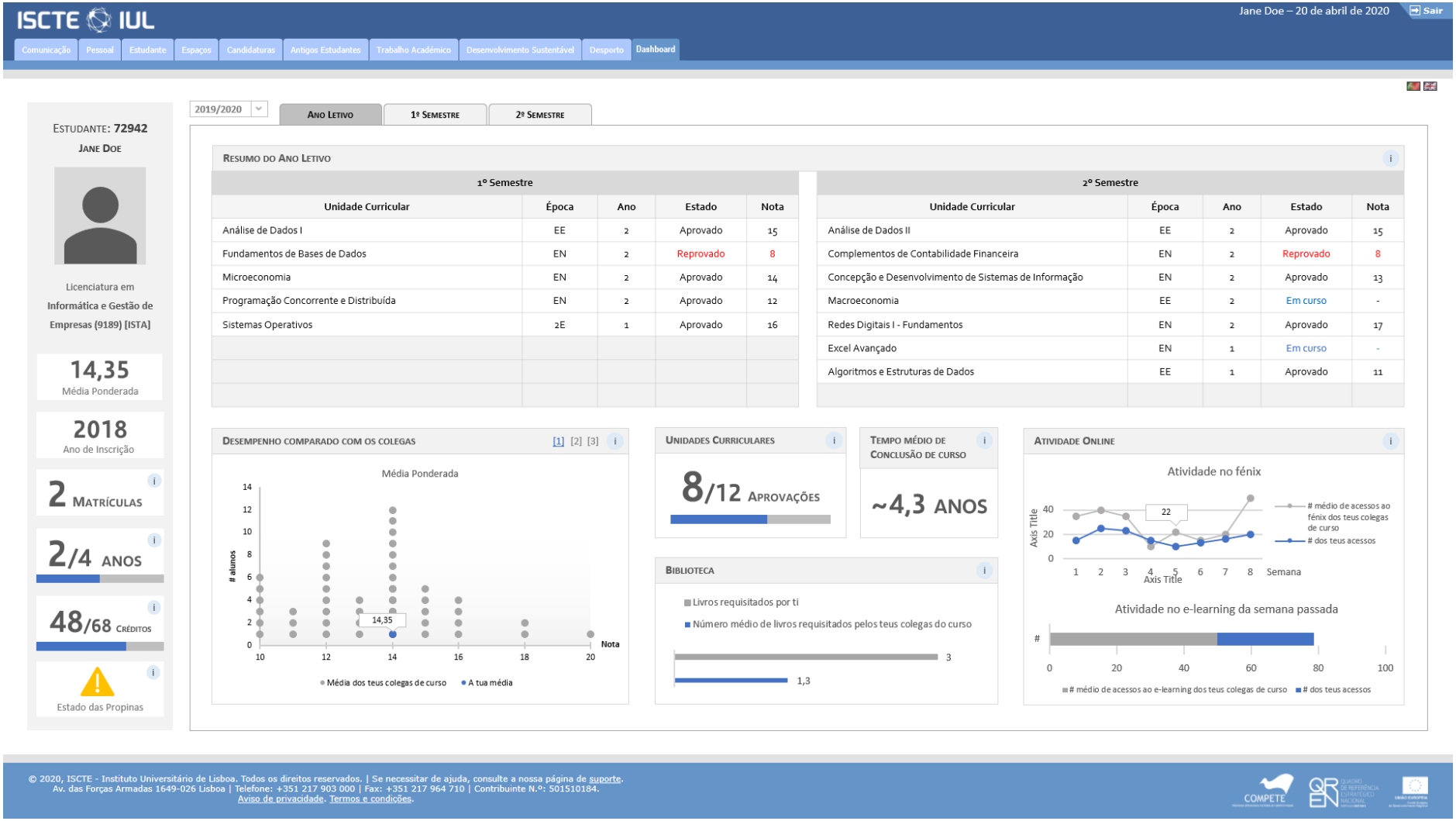

Figure 12: Dashboard prototype view of the academic year with the layout of the Fenix platform (in Portuguese)

Some widgets are accompanied by a help icon ( ) that provides additional information of the data that is being displayed.

Figure 13: Example of a help button.

Moreover, the widget regarding the student's *performance* enables the change of charts in two other options showing the weighted average distribution (chart on the left) and average per course unit (chart on the right) in the following figure.

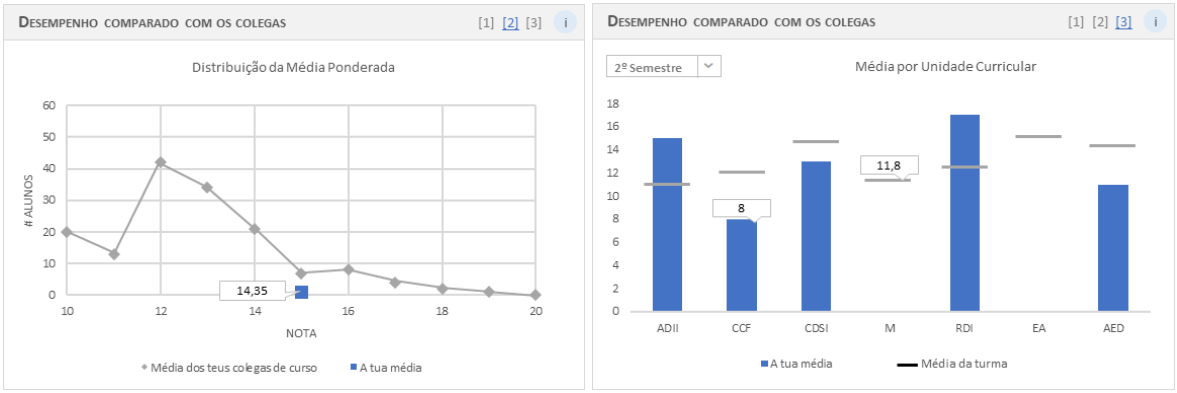

Figure 14: Options [2] and [3] of the feature *Desempenho comparado com os colegas* (Performance compared with peers, in Portuguese)

The following **Table 16** lists the features chosen for the *Academic Year* view and the correspondent region in the *Ano Letivo* tab on the dashboard prototype, represented in **Figure 15**. Additionally, to the features gathered from the literature review, new features were added others in the iterative dashboard design process.

Table 16: Features for the academic year tab

Region	Feature-id	Description
1	1	Personal information (student number and name (first and last), photography)
	2	Programme name
	3	Number of curricular units
	4	Year of first enrolment
	6	Current average programme grade
	7	Credits achieved
	20	Tuition fees status (up to date or late)
	(*)	Degree progress (in years)
2	10	Grades per course
	12	Concluded and failed courses
3	9	Overall student performance compared with peers
4	5	Length of bachelor in years (prediction with historical data)
	8	Number of enrolled courses/Approved courses
	17	Number of times the student accessed the Fénix platform compared with peers
	18	Number of times the student accessed the e-learning platform compared with peers
	19	Library status (number of requisitioned books compared with peers)

(*) new feature

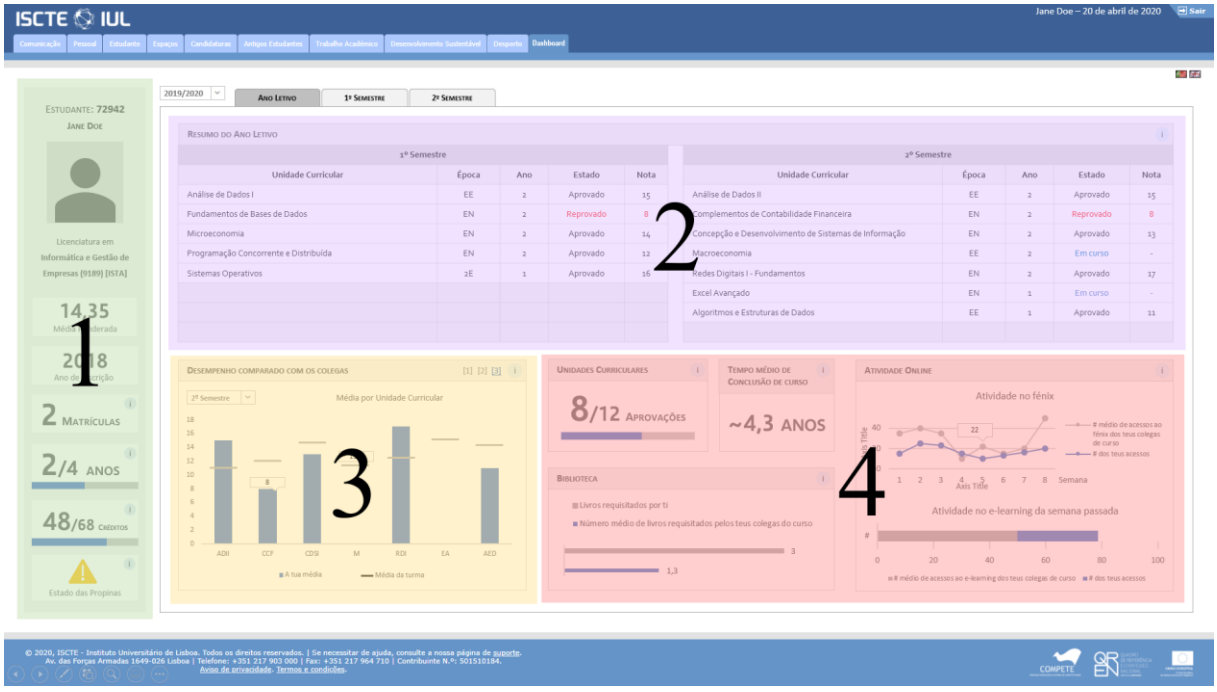

Figure 15: *Ano Letivo* tab by region

3.2. The Semester view

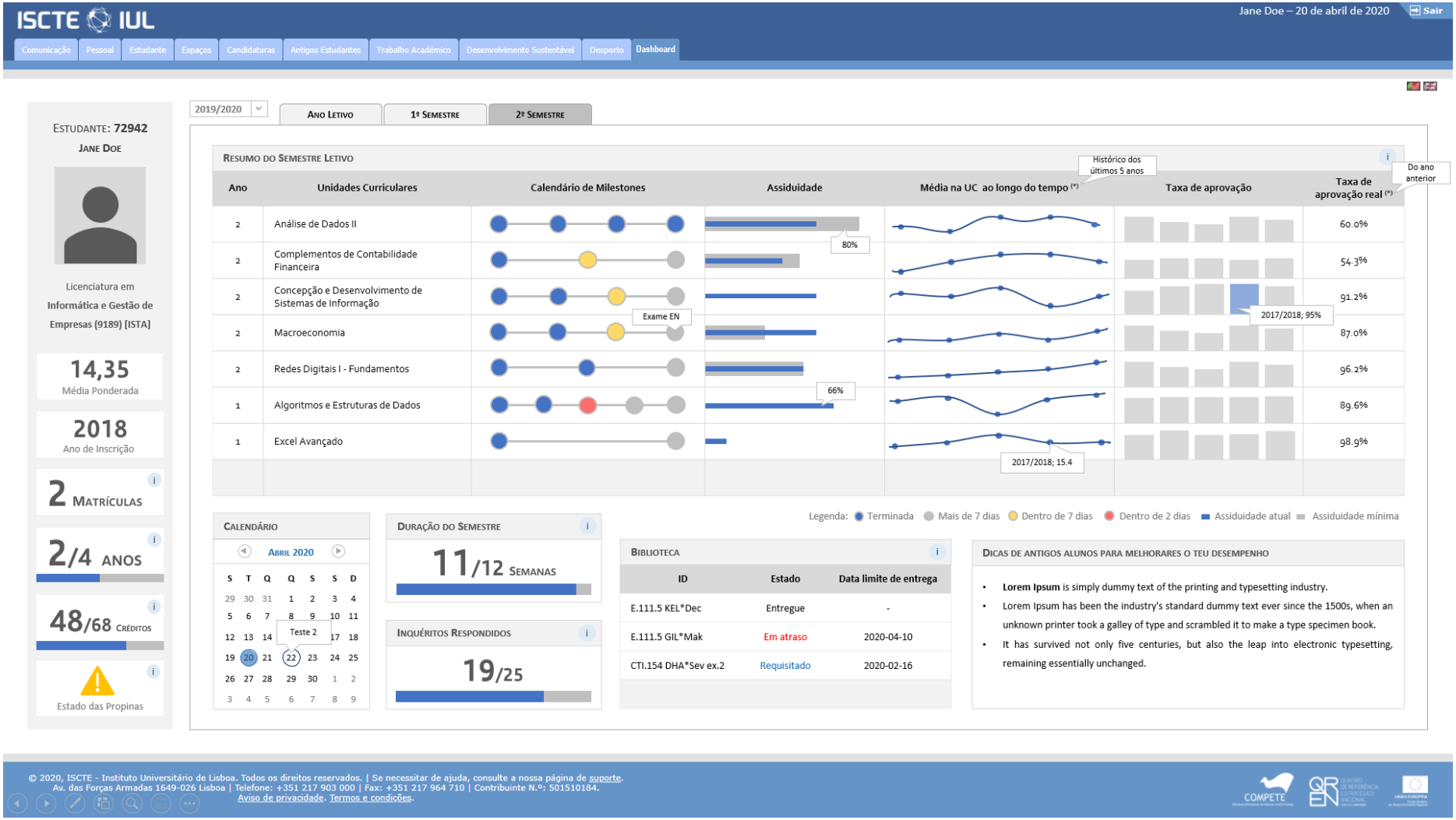

Figure 16: Dashboard prototype view of the 2nd semester with the layout in the Fenix platform (in Portuguese)

The following **Table 17** lists the features chosen for the *semester* view and the correspondent region in the *2º. Semestre* tab on the dashboard proposal, represented in **Figure 17**.

Table 17: Features for the semester tabs

Region	Feature-id	Description
1	1	Personal information
	2	Programme name
	3	Current average programme grade
	4	Year of first enrolment
	6	Number of curricular units
	7	Credits achieved
	20	Tuition fees status (up to date or late)
	(*)	Degree progress in years
2	10	Grades per course
	14	Reminders of upcoming assignments
	15	Attendance (indicating if the student is at risk of failing or not - only applicable for courses with a minimum number of class attendance)
	(*)	Average course grades in the past years
	(*)	Approval rate in the course
3	13	Calendar with evaluation dates
	(*)	Progress in weeks until the end of the semester
	20	Unanswered/answered pedagogical surveys
4	11	Tips as tricks from former students to help students improve their learning study behaviour
	19	Library status (requisitioned books and due dates)

(*) new feature

Figure 17: *2.º Semestre* tab by region

Chapter 4:

DATUS Validation and Results analysis

"Repeated patterns become facts, and when you're in touch with the facts you can predict the future."

(Anieekee Ezekiel)

This chapter concerns the validation of the model proposal developed and its application on the dashboard prototype built for Iscte. Descriptive statistics is used to summarize the data gathered from the case study. Data visualization is used to draw conclusions from the sample.

An exploratory study was undertaken to enable the validation of the DATUS model. This study intended to evaluate the usability of the dashboard proposals (presented in **Chapter 3**) using DATUS and analyse the outcomes of employing the model for the first time.

Three evaluation objectives were identified:

- Assess the applicability of the model on a dashboard for academic purposes, and at the same time evaluate the dimensions and metrics defined
- Retrieve the opinions of the participants on the dashboard proposal;
- Identify improvements for the dashboard prototype.

4.1. Experience construction/design

A user testing session includes different phases [41]:

- Preparation, which includes all the material and environment necessary to receive the participants of the case study;
- Introduction, which is about welcoming the participants and explaining the instructions;
- Testing, which happens when the participants start performing the tasks given by interacting for the first time with the dashboard;
- Redaction, which includes further comments or explanations and the development of a report.

Furthermore, this chapter also describes the planning of the case study.

4.1.1. Development of tasks

Tasks were based on real case situations to simulate how students would interact with the dashboard in a real-life situation. It is essential to include a variety of relevant tasks of different lengths and levels of difficulty to obtain demonstrative and accurate performance. The description of tasks must show the user what to do and not how to do it [41].

The following **Table 18** summarizes the tasks developed for the case study.

Table 18: Translated list of tasks for the case study.

Task-id	Task description
1	Can I change the school year of the information I am viewing on the dashboard?
2	My degree is...
3	My average programme grade is...
4	I've been registered in Iscte since...
5	The year I currently attend is the...
6	How many credits have I made this school year?
7	Regarding the status of tuition fees, are they up to date?
8	How many curricular units are awaiting the grade?
9	How many curricular units have I failed?
10	How many curricular units have I passed?
11	My highest grade was in the curricular units of...
12	In the first semester, how many curricular units am I behind?
13	Can I change the type of performance chart compared to my colleagues?
14	The most frequent programme average among your course colleagues is...
15	On average, how long do students from your course need to finish the course?
16	How many books did you request this school year?
17	The average number of visits to the Fénix of your classmates in week 5 was...
18	Can I consult more detailed information for the second semester?
19	The day I am consulting the dashboard is...
20	How many of last year's curricular units am I attending this academic year?
21	The next milestone belongs to the curricular unit of...
22	The curricular units in which I have already finished all the milestones is...
23	The last milestone of Macroeconomics is...
24	The average grade of Advanced Excel in the year 2017/2018 was...
25	I can consult the average grade of any curricular unit since the school year of...
26	In the Information Systems Design and Development programme, what was the actual approval rate in the 2017/2018 school year?
27	In how many curricular units do I have more than 7 days until the next milestone?
28	How many curricular units am I at risk of failing due to attendance?
29	In the previous school year, 2018/2019, the actual approval rate in the Financial Accounting Complements programme was...
30	Can I find out where I can get more information about the total approval rate?
31	How many weeks do I have before the end of the semester?
32	How many surveys have I answered?
33	Do I have any books requested whose deadline has passed?
34	Can I find the tips from former students?

The original task list, in Portuguese, used in the experiment is available in **Annex D**.

4.1.2. Users, material and agenda

The users must have at least one enrolment in Iscte.

The experiment was done with an interactive version of the dashboard's prototype using Microsoft® PowerPoint®. The studies were done through individual remote meetings of approximately 30 minutes each using *Zoom Video Communications*. The planned agenda was:

- Welcoming the participant (3 minutes)
- Testing the dashboard by performing the given tasks (15 minutes)
- Answering the final questionnaire (10 minutes)
- Closing session and further questions or comments (2 minutes)

The recording started after the participant agreed on recording his screen.

The data collected included the screen recording of the participant, the answers to the list of tasks given and the answers to the final questionnaire, which analysis is made in the following item 4.2.

4.2. Analysis of the results

4.2.1. Descriptive analysis

The experiment was completed with a sample of 30 students from Iscte, from different programmes and cycles of studies. **Figure 18** shows the distribution of the respondents per course cycle (bachelor, master, postgraduate and doctoral). The dashboard was designed for students attending a bachelor's degree, which enlightens the higher demand for undergraduate students (14). Following are students from master's degrees (13) since they have finished their undergraduate degree at Iscte as well, and their input is equally relevant.

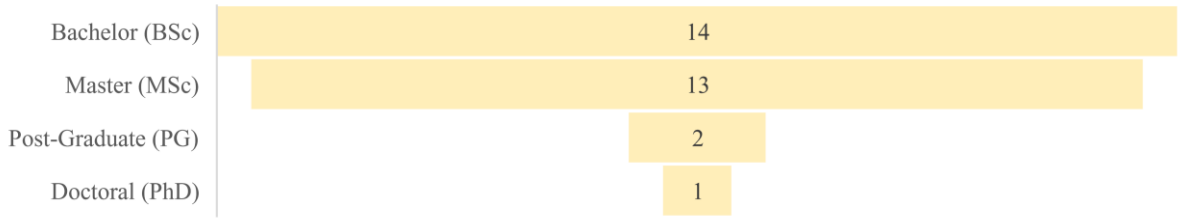

Figure 18: Number of respondents per course cycle

Concerning the number of participants, fourteen students (47%) are attending a bachelor's degree, thirteen students (43%) are attending a master's degree, two students (7%) attended a post-graduate degree, and one student (3%) is attending a doctoral degree but this participant academic background was a bachelor and master's degree from Iscte as well. Furthermore, participants are distributed over 13 different courses, which can be seen in **Figure 19** and the correspondent degree of each course.

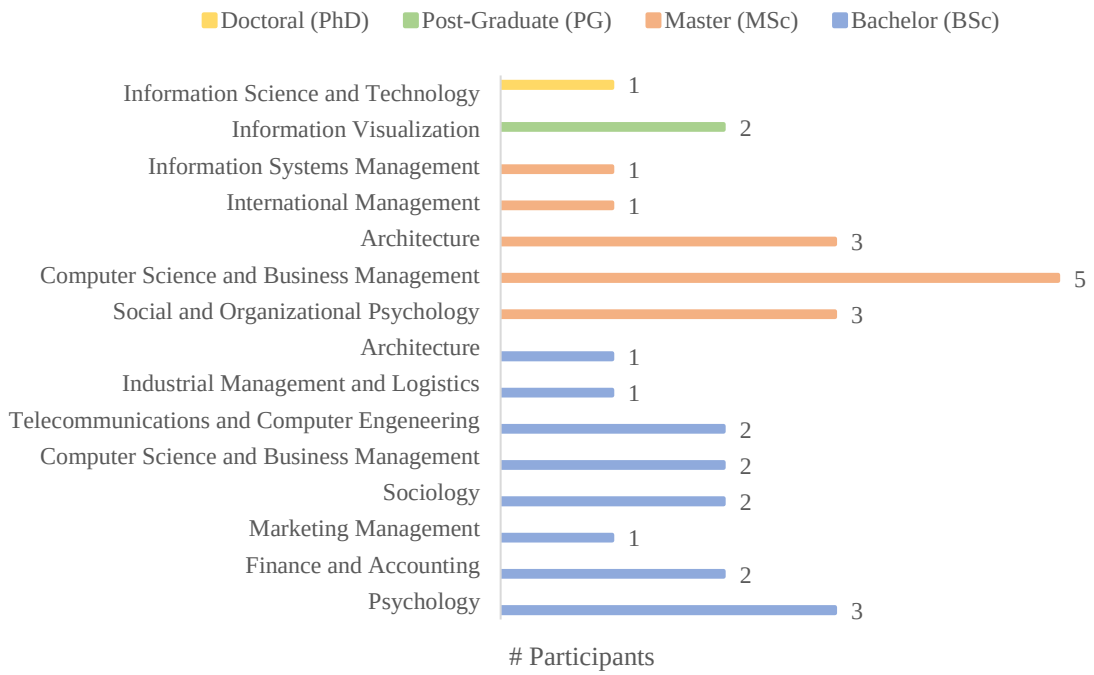

Figure 19: Distribution of courses and degrees of the participants

Table 18 shows the distribution of the participants according to gender. In total, 13 female and 17 male students participated in the case study distributed per the following courses**.**

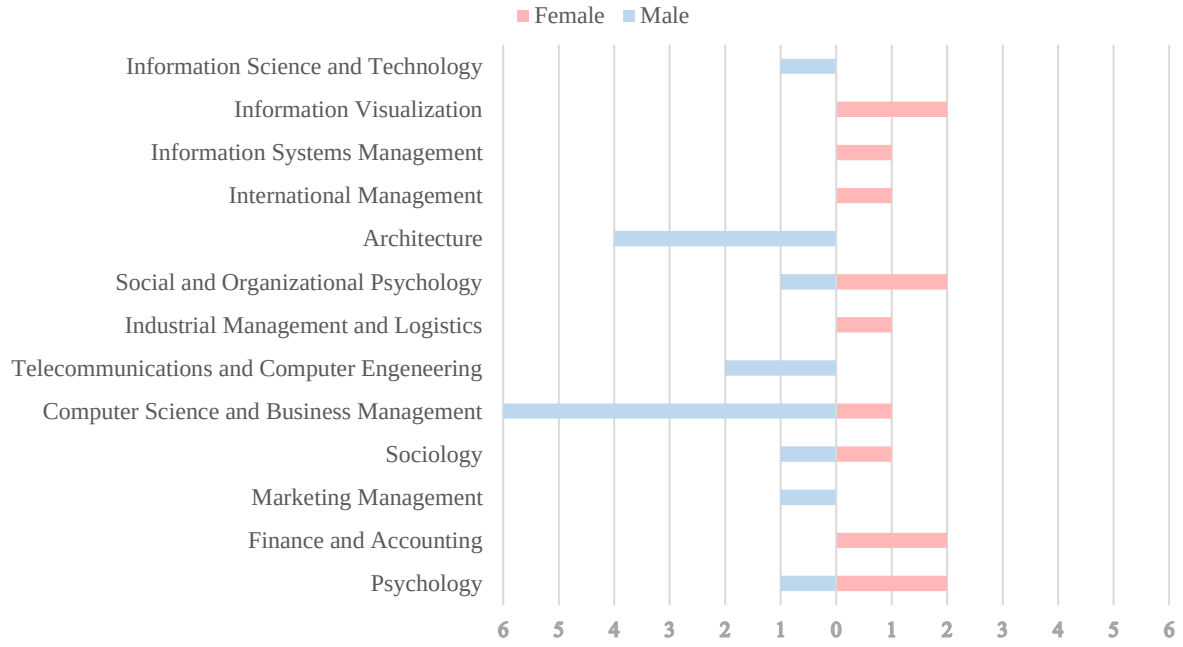

Figure 20: Distribution of participants' gender by courses

To reduce the diversity of the 13 different courses and to help in the discussion of the results, the courses were grouped into three academic or study areas: Social Sciences with eight participants, Technologies and Architecture with 17 participants and Management and Finance with five participants.

4.2.2. DATUS: metrics' analysis

Recalling the proposed model DATUS, for dashboard usability evaluation, some metrics of the model require the collection of information using recording methods. Following, the results of each of those metrics are presented.

Number of analytical goals completed successfully

Reminding the definition of this metric, an analytical goal completed successfully means that each analytical goal must have been finished with the correct outcome.

It is important to note that no participant left blank answers and the few tasks whose answers were marked as «*I don't know*» were not counted as successfully completed but were counted in the next metric as incorrect outcomes. Analysing the data collected from the total of the 30 participants, the following sentences can be outlined:

- At least 82% (28/34) of the tasks were successfully completed by all participants;
- 23% of the participants (7/30) completed all the tasks successfully (34/34);
- 93% of participants (28/30) completed 91% (31/34) of the tasks with a successful outcome;

The distribution of this metric is graphically represented in **Figure 21**.

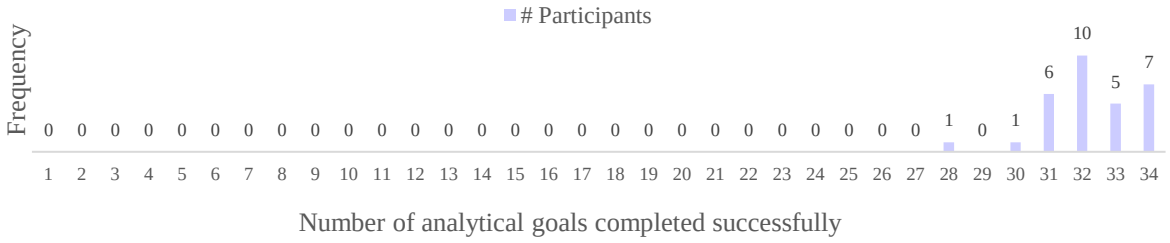

Figure 21: Distribution of the number of analytical goals completed successfully by participants

From the results gathered, the range of analytical goals completed successfully goes from 28 to 34 tasks, and the participants average is 32 tasks completed successfully. The distribution of these values can be seen by gender in the following **Figure 22** and by study areas in **Figure 23**.

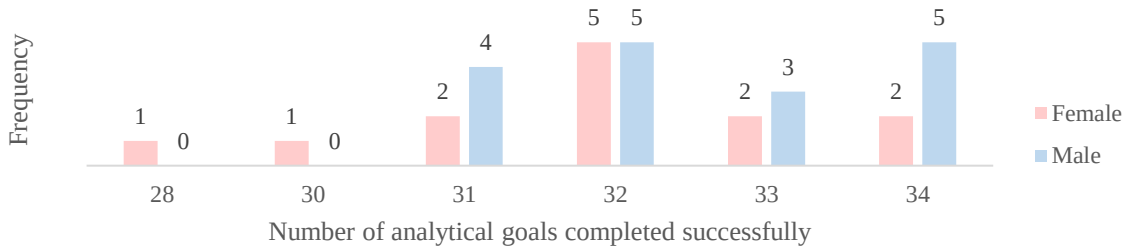

Figure 22: Distribution of the number of analytical goals completed successfully by gender

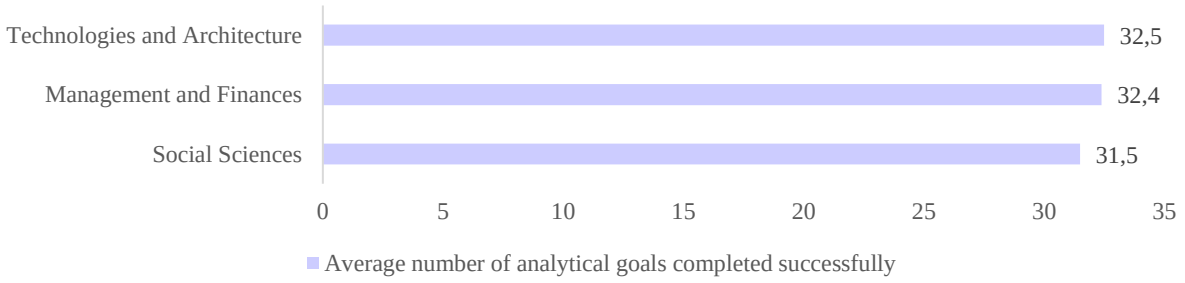

Figure 23: Average number of analytical goals completed successfully by study area

The group with the largest number of complete tasks with the expected outcome is the Technology and Architecture group, which can be explained by the fact that many participants from this group have already attended classes with topics related to data visualization and decision support systems. On average, in this sample, the group of participants from Social Sciences missed one more task with the correct outcome than the other two groups under study. Nonetheless, the difference between groups is not significant. Every group shows promising results in this metric for the usability of the dashboard since it aims to be understood by the whole universe of students with different backgrounds from Iscte.

Number of errors

The definition of this metric in DATUS model is the *average number of incorrect outcomes given by users when trying to achieve a goal*; the goal is a list of tasks in this analysis.

The following figure displays the distribution of participants by number of errors given, i.e., wrong answers in the list of tasks. The range of errors given goes from 0 to 6.

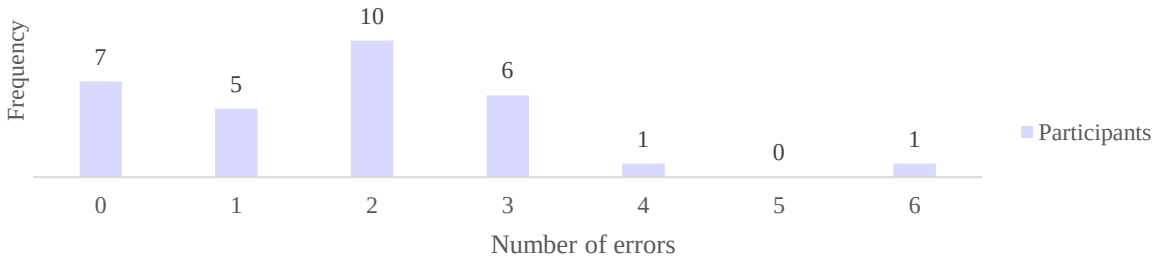

Figure 24: Distribution of the participants by number of errors

About 77% (23/30) of the participants missed the correct outcome of at least one task, meaning that only 23% (7/30) of the participants answered all the questions with the expected result. Two is the most common number of errors given by participants.

This metric is closely related to the previous *number of analytical goals completed successfully*, but the number of errors shows which tasks have led to the most wrong answers. In addition, it is by quantifying these values that one can see where there are the most significant usability problems to be improved in the proposed dashboard.

Table 19: Number of wrong outcomes per task

Task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
# wrong answers	0	0	0	0	0	0	0	1	3	4	0	0	0	1	0	0	3

Task	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	Total
# wrong answers	0	0	7	0	0	4	2	0	3	16	5	2	2	0	0	0	0	53

The 30 participants produced 1020 answers in a list of 34 tasks. The number of incorrect answers was 53 (out of 1020), accounting for an error rate of 5,2%. **Table 19** outlines the tasks with the highest number of errors (i.e., tasks number 10, 20, 23, 27, and 28), which are considered as having four or more wrong answers given.

Table 20: Summary data of the analysis of the number of wrong answers

Number of errors…	Number of tasks	Average of wrong answers
…per total of tasks	34	1,56
	29(*)	0,59
…per user	34	1,77
	29(*)	0,57

(*) total of tasks minus the five tasks with the highest number of wrong answers given

The average number of errors per user is 1.77 wrong answers, as presented in **Table 20**. These averages correspond to answering wrong to 5,2% of tasks, as previously stated. The goal of comparing the average of errors per participant and per task is to understand how much this value varies without the tasks that contribute to this average being higher.

Considering only 29 tasks (i.e., without the five tasks with highest number of errors), the result is 0,59 average wrong answers per total of tasks. These five tasks have an impact of 0.97 percentage points, corresponding to 1 more wrong task per participant. Thus, apart from the tasks with the most errors, if they had been defined more precisely so as to be interpreted by all participants, the average number of wrong tasks per user would have been 0.57 a figure that would help to conclude a better usability by the dashboard.

It is worth to look more closely to these five tasks, and recall what was requested in each of them:

- Task #10: How many course units have I passed?
- Task #20: How many course units am I attending this academic year?
- Task #23: The last milestone of Macroeconomics is…
- Task #27: In how many course units do I have more than seven days until the next milestone?
- Task #28: How many curricular units am I at risk of failing due to attendance?

Further analysis was made of the recorded video of the participants to understand the main reasons to have missed the outcome of these tasks. At the end of the case study session, some participants were

asked about some of their answers, mainly the wrong ones and their answers confirmed the following description on the most common mistakes or misunderstandings among the participants who answered different outcomes from the expected.

Task number 10 **(4 errors)**

- Misreading of data from the dashboard was one of the reasons to answer incorrectly to this task;
- Confusion between tabs, *Ano Letivo* and *2.º Semestre*, due to lack of information in the task description;
- Although they were in the correct tab, they responded only considering information from one semester when the description asked for information about the two semesters;

Task number 20 **(7 errors)**

- Participants were looking at the 2nd Semester tab at the point of answering this task and did not understand that the task referred to the total of curricular units of the academic year;
- Some participants misread the description of the task;

Task number 23 **(4 errors)**

- Some participants gave different answers then the expected, i.e., not wrong but that is not what the description asked for;
- *Last milestone* confused with *next milestone* in task description;

Task number 27 **(16 errors)**

- Lack of attention to reading the task description in its entirety, i.e., answers were given without paying attention to the keyword *next* in the description;
- Incorrect interpretation of the caption of the widget;
- Some participants interpret the caption and the logic behind the widget but counted wrongly;
- After finishing completing all the tasks, some participants were asked if they understood the task and when they read it again many said they did not notice the keyword;

Task number 28 **(5 errors)**

- Some participants read the description wrongly (answered for *which* instead of *how many);*
- Incorrect interpretation of the caption of the widget or failure to find it; the caption may be too small to attract users' attention.

These detailed explanatory actions also summarise what was observed in the other tasks with a negligible number of incorrect responses. Despite these remarks, there may have been a lack of clarity

in the description of the tasks otherwise the largest number of errors given were not concentrated in these five tasks, which represent about 15% of tasks.

The values obtained from these two metrics, analytical goals completed successfully and the number of errors, contribute positively to the usability assessment of the dashboard proposal regarding effectiveness aspects.

Number of actions to complete a task

The number of actions to complete a task was defined to be measured as the *number of clicks to complete a task* in the model proposal. Nonetheless, given the difference each task needs, some questions require a yes/no answer, others need to write a more extensive answer, making the tasks very different and uneven in terms of number of interactions with the dashboard. Consequently, it requires too laborious analysis to analyse the number of clicks for each task. Furthermore, users have not been instructed to restrict themselves to the completeness of tasks, but rather to exploit the dashboard according to their own needs and curiosity. Considering that the dashboard is an exploratory interface, the case study revealed that many participants indeed clicked to explore what the dashboard could offer and not narrowly to answer the tasks given.

It was considered to redefine this metric. Averaging the results is not an option, since the average number of clicks per task is not the most appropriate metric for assessing dashboard efficiency. Thus, after analysing the cost of obtaining information for the analysis in the context of a dashboard, it was decided to discard this metric. The efficiency dimension in the DATUS model can still be evaluated using the remaining metrics (see **Table 12** in **Chapter 2**).

Task time of the first use

This metric is about the *total time, in seconds, for users to execute a given list of tasks (or analytical goal) for the first time* they consult the dashboard.

Table 21 summarizes two variables to be considered in this metric about how long it would take for a student to use the broadest range of functions available on the dashboard for the first time. The dashboard was given to an expert, which took about five minutes to perform all the tasks and answer all the questions. Therefore, the planned time for the participants to do the same exercise should be around 15 minutes, three times more the expert's time.

Despite the firm attempt to count the total useful time of the participants in the exercise to perform all tasks, some recordings were influenced with minimal breaks caused by technical problems and/or questions not related to the answers of the tasks.

Table 21: Time of the first use to complete the list of tasks

User-id	Total time		Average time per task
	Minutes	Seconds	Seconds (aprox.)
1	18	55	33
2	16	29	29
3	15	31	27
4	13	37	24
5	12	44	22
6	19	47	35
7	18	39	33
8	18	6	32
9	15	29	27
10	22	47	40
11	13	53	25
12	13	41	24
13	16	38	29
14	13	2	23
15	11	53	21
16	12	33	22
17	16	33	29
18	16	16	29
19	7	55	14
20	22	48	40
21	12	30	22
22	22	29	40
23	12	20	22
24	10	30	19
25	12	46	23
26	9	43	17
27	13	29	24
28	17	8	30
29	20	17	36
30	14	3	25
AVERAGE	**15**	**21**	**27**

The 30 participants performed all the tasks with different times within a range from 7 minutes and 55 seconds to 22 minutes and 48 seconds. From this, the average time to complete the list of tasks was 15 minutes and 21 seconds, consisting of an average of 27 seconds per task. The positive point we can outline from the analysis of this metric is the match between the total time previously planned (15 minutes) for participants to perform the given tasks for the first time and the actual average time (15 minutes and 21 seconds) of this sample of participants that completed the tasks.

The following histogram, in **Figure 25**, enables the analysis of the distribution of the number of participants by time intervals in minutes. Half of the participants (15/30) are placed between 12 and 17 minutes for the task's completion, 11 participants took more than 17 minutes to perform all tasks and

four participants took between 8 and 12 minutes to complete the overall analytical goal, i.e., the list of tasks.

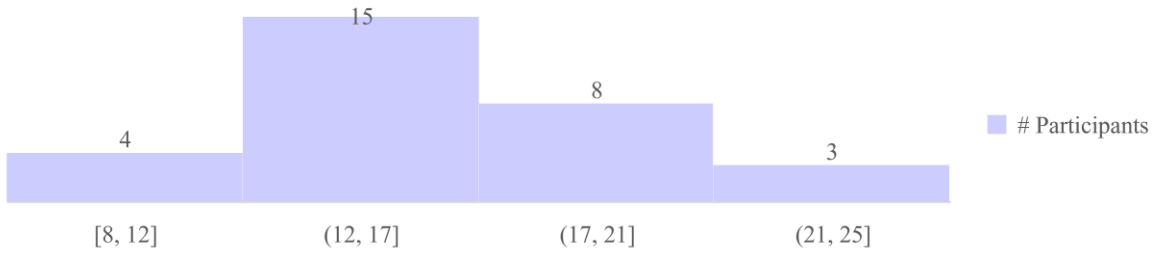

Figure 25: Histogram of the total time of the participants, in intervals of whole minutes

Figure 26 shows the average total time to complete the list of tasks per study group. As can be seen, Management and Finance and Social Sciences have similar time, differing by only 0,2 minutes. As for Technologies and Architecture, it has the best average time. However, the differences between the groups are not significant as the it is only in seconds and never more than a minute apart. These values are consistent with the general average of approximately 15 minutes.

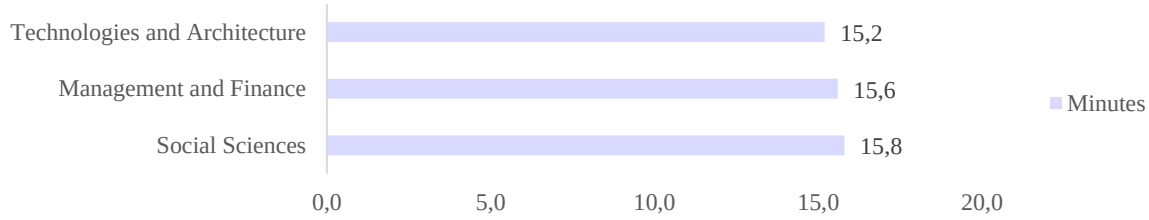

Figure 26: Time to complete all tasks per study area

Task time of the second use

Even though it is present in the proposed model, due to limitations of time in this dissertation, it could not be tested.

Current date and date of data extraction/Current or date of data extraction/None

This metric corresponds to a technical verification regarding the dashboard information, it represents an indirect metric regarding data trustfulness.

The prototype was designed considering the look and feel of the Fénix platform, and the platform itself always displays the current date of when it is being accessed, placed in the top right corner of the web page. Due to this information, it was thought that the data on the dashboard should be updated on the day and time that the dashboard is being consulted, or the Dashboard tab is being accessed.

As a consequence of the tests and feedback gathered, the time of the last update will be added to the dashboards proposal, as students indicated that they did not know the time of the last data update (which is significant, for instance when they are waiting for a grade to come out and it could happen at a specific time).

Comments and suggestions for the dashboard

The comments and suggestions received from the participants are compiled in **Annex D**.

4.2.3. DATUS: satisfaction questionnaire analysis

As defined in the DATUS model, the analysis of the questionnaire is done considering the rules of the questionnaires used for the construction of the DATUS questionnaire. Those rules are summarised in the following table.

Table 22: Rules for calculating DATUS scores

Score name	Average the responses to:
Satisfaction	Items 1, 6, 16 and 18
Effectiveness	Items 3 and 13
Efficiency	Items 4 and 5
Operability	Items 2, 9, 11, 14 and 20
Learnability	Items 7, 8, 10, 12
User interface aesthetics	Item 15
Appropriate recognizability	Item 17
Accessibility	Item 19

In order to apply the previous rules to evaluate the respective dimensions under study, the data for the calculation of the scores were summarised.

Table 23: Summary data of DATUS questionnaire by question

Dimension	Question-id	Mean	Median	Standard Deviation
Accessibility	**19**	4,6	5	0,498
Appropriate recognizability	**17**	4,76	5	0,511
Effectiveness	**3**	4,43	5	0,679
	13	4,67	5	0,547
Efficiency	**4**	4,5	5	0,682
	5	4,7	5	0,535
Learnability	**7**	4,7	5	0,466
	8	4,47	5	0,681
	10	4,8	5	0,407
	12	4,67	5	0,479
Operability	**2**	4,6	5	0,498
	9	4,43	5	0,774
	11	4,33	4	0,547
	14	4,63	5	0,557
	20	4,73	5	0,521
Satisfaction	**1**	4,7	5	0,467
	6	4,67	5	0,479
	16	4,77	5	0,430
	18	4,83	5	0,379
User interface aesthetics	**15**	4,73	5	0,521
	Score	92,73	99	-

The dashboard prototype has a mean score of 92,73 out of 100 and a median score of 99.

Table 24: Summary data of DATUS questionnaire by dimension

Dimension	Questions	Mean	Median	Standard Deviation
Accessibility	19	4,6	5	0,498
Appropriate recognizability	17	4,76	5	0,511
Effectiveness	3, 13	4,55	5	0,613
Efficiency	4, 5	4,6	5	0,609
Learnability	7, 8, 10, 12	4,66	5	0,508
Operability	2, 9, 11, 14, 20	4,55	4,8	0,579
Satisfaction	1, 6, 16, 18	4,74	5	0,439
User interface aesthetics	15	4,73	5	0,521

Appropriate recognizability and *satisfaction* are the usability dimensions with the best results, 4,76 and 4,74, respectively and still with a very similar value of 4,73, *user interface aesthetics*. Although not indicating values below 4, the dimensions with lower averages are *effectiveness* and *operability* with a mean of 4,55 each. Regarding *operability*, this result may have been influenced by the fact that the dashboard is not a final and fully developed solution and participants' perception may have been influenced by technical flaws because it is a prototype.

Figure 27 supports the previous **Table 24** and translates the association between dimensions under study, and the averages got from the answers gathered from the questionnaire.

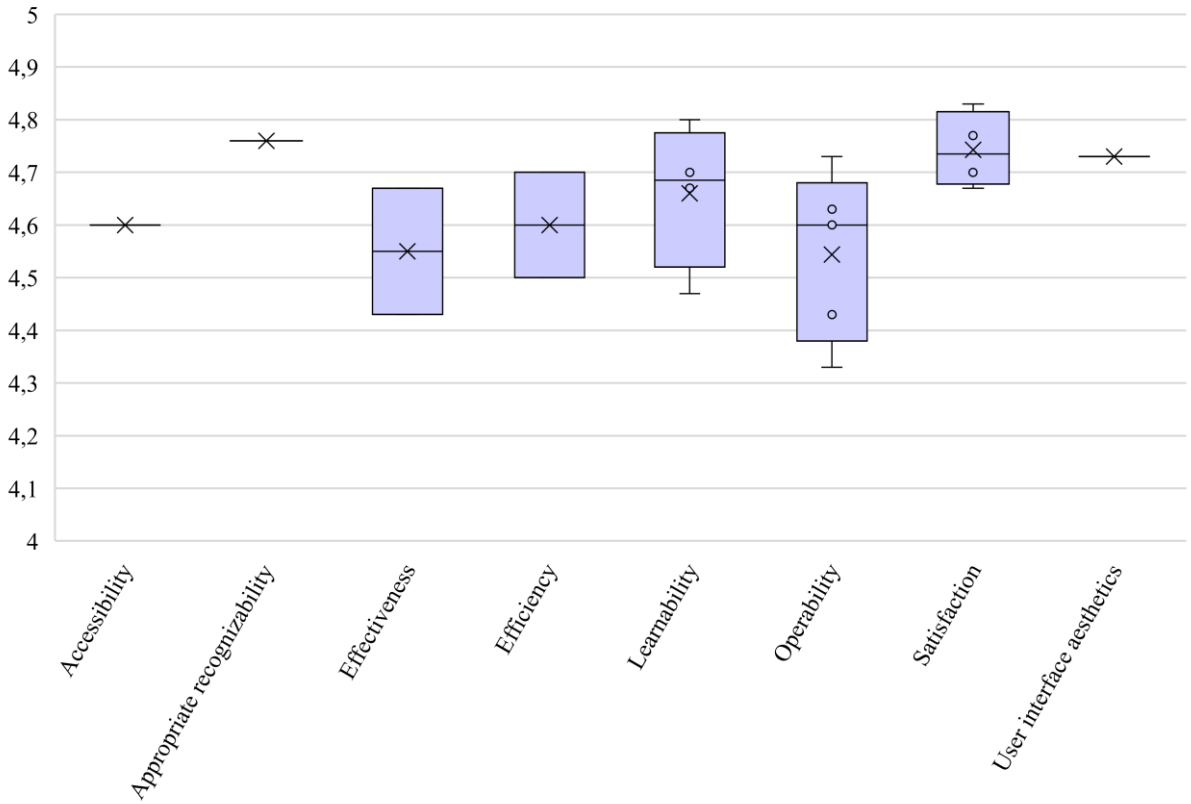

Figure 27: Distribution of answers from DATUS questionnaire by dimension

The X-axis was adapted to starting at the value four from the Likert-scale, as noted above because there is no average value below four. The boxplot of DATUS questionnaire data shows higher scores in *satisfaction* and *appropriate recognizability* with the dashboard prototype.

Accessibility, *appropriate recognizability* and *user interface aesthetics* are measured from a single question in the questionnaire, which explains its formats in the figure. Overall participants have a high level of agreement with each other regarding *effectiveness* and *efficiency*, following a normal distribution, although these dimensions have the lowest averages.

No outliers have been identified in the responses to the questionnaire.

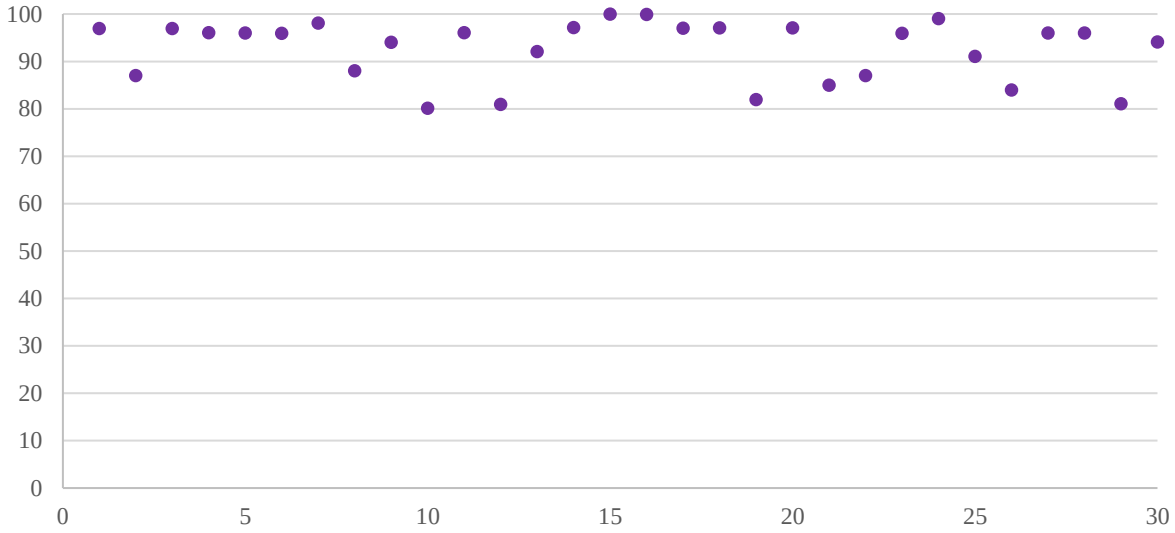

Figure 28: Scores per participant

As defined in the model proposal, scores are ways to measure the different dimensions in the questionnaire and the closer to 100 (maximum value achievable), the better. **Figure 28** illustrates the overall score given by the 30 participants, regardless the dimensions. The lowest score was 80 and the highest was 100. In general, the results were quite favourable due to the proximity of the dots to the maximum value.

Conclusion

"It's the job that's never started that takes longest to finish."
(J. R. R. Tolkien)

Today many products are launched; however, organizations are not always able to prevail in the competitive software product market. Usability is an essential part of the software development, and the right investment brings several benefits; such as more streamlined products, better user experience with their product and increase of customer's conversion; in fact, usability is not only good for users but for business, enabling companies to save time during the development process and reduce training [86]. The study of usability in products is a differentiating factor, concerning the user experience and the care in creating something to satisfy several dimensions of interest to the attract and maintain users and not only to solve a problem of one dimension only.

Short and quick usages of a software product are common recurrence. When a user encounters any difficulty in using or understanding a product, the first reaction is to leave and search for a substitute among the competition. With so much supply, users prefer to waste the minimum time finding what they are looking for rather than dealing with an interface that is poor in ease of interaction. In the context of dashboards, whether the interface fails to clearly state what it can offer or is difficult to read the essential information, the interest of users will decrease and consequently their engagement with the tool.

This dissertation addressed a relevant topic to contribute to the scientific community. This dissertation aimed to develop two artefacts: a usability measurement model and a method in order to help dashboard designers to identify usability problems and make improvements on their interfaces. Regarding the artefacts, a dashboard for the Fénix platform was also developed for Iscte to serve as a contribution and useful tool for students to monitor their performance throughout their academic career and possible future implementation.

The literature review significantly supported both the model and the dashboard proposal. Hence, it was found that most of the dimensions measured in other studies with different products met the dimensions defined by the International Organization for Standardization regarding usability and quality. Thus, all dimensions of the proposed model DATUS have the national standards from ISO/IEC 9241-11:2018 and ISO/IEC 25010:2011 as the primary source. DATUS includes eight dimensions of which corresponds to a specific facet of usability. One or more metrics accompanies each dimension. Moreover, to complete the model, two methods were defined to measure usability, not only because they are the most used among usability studies but also because the recording method allows a more careful later analysis and the questionnaire is easily administrative. Except for one metric that turned out to be unsuitable for the context of dashboards, all others proved to be feasible for measuring the usability of these types of interfaces.

Regarding the student dashboard design, investing in the implementation of a dashboard for Iscte would be something very well received by the students; this is sustained not only by the statistical results regarding their satisfaction but also by the feedback during the tests performed. Most of the time, the participants have voluntarily expressed positive comments about their first impression on the dashboards. Translating the case study into numbers and observations, 93% of participants completed 91% of tasks successfully, and 23% completed all tasks with the expected outcome. There was a 5.2% error rate, and the main reasons for errors in participants' responses were a lack of attention or clarity in the task description. Participants took the expected time to complete the tasks defined for the study. There were no significant differences in task completion between the groups defined by study area. Furthermore, the final analysis showed promising results in terms of the usability of the dashboard, yet a handy list of comments and suggestions for improvement was collected.

We may conclude that students could benefit with the use of a dashboard to monitor their performance throughout the academic year and make decisions based on relevant data for their academic career, minimizing the possible lack of engagement from students.

Regarding the research question: *how to measure the usability of dashboards, quickly and effectively?*, the current work has achieved the contributions outlined in the Introduction. A model and an accompanying method for *measuring the usability of dashboards* were developed. DATUS made it possible to effectively check the usability of a dashboard, fulfilling the objectives outlined at the beginning of the dissertation, and providing dashboard designers with support in identifying usability issues. Typically, dashboards designers are not the same people who will use the dashboards. Moreover, their understanding of the dashboard's functionalities and usability may be different from the end-users' perspective. The application of the DATUS model enhances the knowledge of the real needs and perceptions of end-users, which is an essential activity of usability measurement. The proposed method proved to be *quick and effective* in evaluating the usability of the dashboard on which it was applied. Nonetheless, it may still be possible to optimize the evaluation method and improve the model due to the limitations described below.

Limitations

A few limitations were identified in the course of this dissertation. First, the applicability of the model in a single case and dashboard. This limited the conclusions about the assessment of the adaptability of DATUS and usefulness to a unique environment.

Secondly, the time restriction did not allow the measurement of the *task time of the second use* of the dashboard. This limited the assessment of the *learnability* dimension to two out of three of its defined metrics. However, this impact was not significant since it was possible to measure this dimension through the other two metrics.

Due to time restrictions, a reliability study of the specific DATUS questionnaire was not performed. The adopted solution consisted in an adaptation of two well-known and studied questionnaires - CSUQ and QUIS. This solution was found reasonable given the existing evidence of construct validity and reliability levels of these two questionnaires.

A final limitation was the usage of a prototype dashboard, instead of an already deployed solution. To mitigate this limitation, the look and feel, as well as the interaction modes in the prototype were made as close as possible to a real dashboard. However, filters were not fully operationalized in the prototype.

Future work

Concerning future work, there are still improvements to be made both in the model and in the student dashboard design.

The development of a second version of the model could include two additional dimensions considered significant from the literature review: *safety* and *memorability*.

We also intent to use and apply the DATUS model in a real project context, with Business Intelligence consultants and dashboard developers.

As far as the Iscte student dashboard is concerned, a new experience can be done with a larger sample of students and an improved second version of the dashboard. The new version of the student dashboard should take into account the comments and suggestions collected from the participants' tests since they are effectively Iscte students and shared what they really think is essential for a better experience when monitoring their performance.

References

[1] D. Norman, “The Design of Everyday Things,” *Des. Everyday Things*, 2016, doi: 10.15358/9783800648108.

[2] J. Gulliksen and I. Boivie, “The Usability Design Process – Integrating User-centered Systems Design in the Software,” vol. 131, pp. 111–131, 2003, doi: 10.1002/spip.174.

[3] A. Seffah, M. Donyaee, R. B. Kline, and H. K. Padda, “Usability measurement and metrics: A consolidated model,” *Softw. Qual. J.*, vol. 14, no. 2, pp. 159–178, 2006, doi: 10.1007/s11219-006-7600-8.

[4] Nigel Bevan, “Usability As Quality of Use,” *Softw. Qual. J.*, vol. 130, pp. 115–130, 1995.

[5] N. Bevan and M. Macleod, “Usability measurement in context,” *Behav. Inf. Technol.*, vol. 13, no. 1–2, pp. 132–145, 1994, doi: 10.1080/01449299408914592.

[6] J. Baumgartner, A. Sonderegger, and J. Sauer, “No need to read: Developing a pictorial single-item scale for measuring perceived usability,” *Int. J. Hum. Comput. Stud.*, vol. 122, no. July 2018, pp. 78–89, 2019, doi: 10.1016/j.ijhcs.2018.08.008.

[7] A. Fernandez, E. Insfran, and S. Abrahão, “Usability evaluation methods for the web: A systematic mapping study,” *Inf. Softw. Technol.*, vol. 53, no. 8, pp. 789–817, 2011, doi: 10.1016/j.infsof.2011.02.007.

[8] A. W. Astin, “Student involvement: A developmental theory for higher education,” *Coll. Student Dev. Acad. Life Psychol. Intellectual, Soc. Moral Issues*, no. July, pp. 251–263, 2014.

[9] L. R. Halverson and C. R. Graham, “Learner engagement in blended learning environments: A conceptual framework,” *Online Learn. J.*, vol. 23, no. 2, pp. 145–178, 2019, doi: 10.24059/olj.v23i2.1481.

[10] J. P. Campbell, P. B. DeBlois, and D. G. Oblinger, “Academic Analytics: A New Tool for a New Era,” *Educ. Rev.*, vol. 42, no. 4, pp. 40–57, 2007, [Online]. Available: https://er.educause.edu/articles/2007/7/academic-analytics-a-new-tool-for-a-new-era.

[11] A. R. Hevner, S. T. March, J. Park, and S. Ram, “DESIGN SCIENCE IN INFORMATION SYSTEMS RESEARCH,” *AI Soc.*, vol. 28, no. 1, pp. 75–105, 2004, doi: 10.1007/BF01205282.

[12] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A Design Science Research Methodology for Information Systems Research A Design Science Research Methodology for Information Systems Research,” vol. 1222, 2014, doi: 10.2753/MIS0742-1222240302.

[13] H. Gioti, S. T. Ponis, and N. Panayiotou, “Social business intelligence: Review and research directions,” *J. Intell. Stud. Bus.*, vol. 8, no. 2, pp. 23–42, 2018.

[14] V. H. Trieu, “Getting value from Business Intelligence systems: A review and research agenda,” *Decis. Support Syst.*, vol. 93, pp. 111–124, 2017, doi: 10.1016/j.dss.2016.09.019.

[15] K. Kawamura and A. C. Schultz, “OVERVIEW OF BUSINESS INTELLIGENCE MATURITY MODELS,” *Proc. - IEEE Int. Work. Robot Hum. Interact. Commun.*, vol. 2005, 2005, doi: 10.1109/ROMAN.2005.1513741.

[16] A. Shollo and K. Kautz, “Association for Information Systems AIS Electronic Library (AISeL) Towards an Understanding of Business Intelligence Towards an Understanding of Business Intelligence,” 2010, [Online]. Available: http://aisel.aisnet.org/acis2010%5Cnhttp://aisel.aisnet.org/acis2010/86.

[17] S. Rouhani, S. Asgari, and S. Vahid Mirhosseini, “Review Study: Business Intelligence Concepts and Approaches,” *Am. J. Sci. Res. ISSN*, vol. 1450, no. 50, pp. 62–75, 2012, [Online]. Available: http://www.eurojournals.com/ajsr.htm.

[18] F. Burstein and C. W. Holsapple, *Handbook on Decision Support Systems 2: Variations*, vol. 1, no. 9. Springer-Verlag Berlin Heidelberg, 2008.

[19] A. Frederiksen, *Competing on analytics: The new science of winning*, vol. 20, no. 5. 2009.

[20] V. C. S. Hsinchun Chen, Roger H. L. Chiang, “Business intelligence and analytics: From big data to big impact,” *MIS Q. Manag. Inf. Syst.*, vol. 36, no. 4, pp. 1165–1188, 2012, doi: 10.5121/ijdps.2017.8101.

[21] P. Long and G. Siemens, “Penetrating the Fog: Analytics in Learning and Education,” *Educ. Rev.*, vol. 46, no. 5, pp. 30–32, 2011, [Online]. Available:

http://search.proquest.com.proxy.library.vanderbilt.edu/docview/964183308/13AF5BC47C138E29FF2/5?accountid=14816%5Cnhttps://login.proxy.library.vanderbilt.edu/login?url=http://search.proquest.com/docview/964183308/13AF5BC47C138E29FF2/5?accountid=14816.
[22] E. P. Lim, H. Chen, and G. Chen, "Business intelligence and analytics: Research directions," *ACM Trans. Manag. Inf. Syst.*, vol. 3, no. 4, pp. 1–10, 2013, doi: 10.1145/2407740.2407741.
[23] C. A. Anzanello, "OLAP Conceitos e Utilização," *Univ. Fed. do Rio Gd. do Sul*, no. Dm, p. 7, 2008.
[24] J. M. Loureiro Torres, "Tendências de Business Intelligence: SSBI como foco principal de estudo," pp. 1–76, 2016.
[25] A. Maseleno, N. Sabani, M. Huda, R. Ahmad, K. Azmi Jasmi, and B. Basiron, "Demystifying Learning Analytics in Personalised Learning," *Int. J. Eng. Technol.*, vol. 7, no. 3, p. 1124, 2018, doi: 10.14419/ijet.v7i3.9789.
[26] T. Elias, "Learning Analytics : Definitions , Processes and Potential," 2011.
[27] P. J. Goldstein and R. N. Katz, "Academic Analytics : The Uses Of Management Information And Technology In Higher Education," *Educ. Q.*, vol. 8, no. December, p. 113, 2005, doi: 10.1080/17439880802097659.
[28] M. A. Chatti, A. L. Dyckhoff, U. Schroeder, and H. Thüs, "A reference model for learning analytics," *Int. J. Technol. Enhanc. Learn.*, vol. 4, no. 5–6, pp. 318–331, 2012, doi: 10.1504/IJTEL.2012.051815.
[29] R. Ferguson, "Learning analytics: Drivers, developments and challenges," *Int. J. Technol. Enhanc. Learn.*, vol. 4, no. 5–6, pp. 304–317, 2012, doi: 10.1504/IJTEL.2012.051816.
[30] K. Withee, *Microsoft Business Intelligence For Dummies*. 2010.
[31] D. Santiago Rivera and G. Shanks, "A Dashboard to Support Management of Business Analytics Capabilities," *J. Decis. Syst.*, vol. 24, no. 1, pp. 73–86, 2015, doi: 10.1080/12460125.2015.994335.
[32] Y. Park and I. H. Jo, "Development of the learning analytics dashboard to support students' learning performance," *J. Univers. Comput. Sci.*, vol. 21, no. 1, pp. 110–133, 2015.
[33] N. R. Aljohani, A. Daud, R. A. Abbasi, J. S. Alowibdi, M. Basheri, and M. A. Aslam, "An integrated framework for course adapted student learning analytics dashboard," *Comput. Human Behav.*, vol. 92, pp. 679–690, 2019, doi: 10.1016/j.chb.2018.03.035.
[34] J. Kim, I. H. Jo, and Y. Park, "Effects of learning analytics dashboard: analyzing the relations among dashboard utilization, satisfaction, and learning achievement," *Asia Pacific Educ. Rev.*, vol. 17, no. 1, pp. 13–24, 2016, doi: 10.1007/s12564-015-9403-8.
[35] K. Verbert, E. Duval, J. Klerkx, S. Govaerts, and J. L. Santos, "Learning Analytics Dashboard Applications," *Am. Behav. Sci.*, vol. 57, no. 10, pp. 1500–1509, 2013, doi: 10.1177/0002764213479363.
[36] T. Broos, M. Pinxten, M. Delporte, K. Verbert, and T. De Laet, "Learning dashboards at scale: early warning and overall first year experience," *Assess. Eval. High. Educ.*, vol. 0, no. 0, pp. 1–20, 2019, doi: 10.1080/02602938.2019.1689546.
[37] L. Lim, S. Joksimović, S. Dawson, and D. Gašević, "Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference?," *ACM Int. Conf. Proceeding Ser.*, pp. 250–259, 2019, doi: 10.1145/3303772.3303804.
[38] S. Charleer, A. Vande Moere, J. Klerkx, K. Verbert, and T. De Laet, "Learning Analytics Dashboards to Support Adviser-Student Dialogue," *IEEE Trans. Learn. Technol.*, vol. 11, no. 3, pp. 389–399, 2018, doi: 10.1109/TLT.2017.2720670.
[39] I. Arriaran Olalde and N. Ipiña Larrañaga, "Indicadores necesarios para diseñar un dashboard desde la perspectiva de los profesores: un estudio cualitativo," *Rev. Iberoam. Educ.*, vol. 80, no. 1, pp. 157–166, 2019, doi: 10.35362/rie8013462.
[40] C. Gong, "Human-computer interaction: The usability test methods and design principles in the human-computer interface design," *Proc. - 2009 2nd IEEE Int. Conf. Comput. Sci. Inf. Technol. ICCSIT 2009*, pp. 283–285, 2009, doi: 10.1109/ICCSIT.2009.5234724.
[41] D. Gonçalves, M. Fonseca, and P. Campos, *Introdução ao Design de Interfaces - Desenho de Ecrãs*, 2nd ed. 2017.
[42] D. Dowding and J. A. Merrill, "The Development of Heuristics for Evaluation of Dashboard Visualizations," *Appl. Clin. Inform.*, vol. 9, no. 3, pp. 511–518, 2018, doi: 10.1055/s-0038-

1666842.
[43] J. Nielson, “Jakob Nielsen ’ s Ten Usability Heuristics,” *Nielsen Norman Gr.*, p. 1, 1995, [Online]. Available: https://tfa.stanford.edu/download/TenUsabilityHeuristics.pdf.
[44] J. Juran and A. Godfrey, *JURAN’S QUALITY HANDBOOK, 5th EDITION*. 1998.
[45] A. D. Garvin, “What Does ‘Product Quality’ Really Mean?,” *Sloan Management review*. pp. 25–43, 1984.
[46] I. E. Commission and I. O. for Standardization, “Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models,” vol. 2010, p. 44, 2010, [Online]. Available: papers2://publication/uuid/4E7AB6E6-6F61-4D11-A7BA-9644832A7308.
[47] J. Nielsen, “Usability 101 : Introduction to Usability Why Usability is Important,” *Usabilty 101introduction to usabilty*, vol. 266, p. 9/2, 2012.
[48] A. Abran, A. Khelifi, W. Suryn, and A. Seffah, “Usability meanings and interpretations in ISO standards,” *Softw. Qual. J.*, vol. 11, no. 4, pp. 325–338, 2003, doi: 10.1023/A:1025869312943.
[49] A. Abran, A. Khelifi, W. Suryn, and A. Seffah, “Consolidating the ISO usability models,” *Proc. 11th Int. Softw. Qual. Manag. Conf.*, pp. 23–25, 2003.
[50] M. Speicher, “What is Usability? A Characterization based on ISO 9241-11 and ISO/IEC 25010,” pp. 1–10, 2015, [Online]. Available: http://arxiv.org/abs/1502.06792.
[51] J. R. Lewis, “Measuring Perceived Usability: SUS, UMUX, and CSUQ Ratings for Four Everyday Products,” *Int. J. Hum. Comput. Interact.*, vol. 35, no. 15, pp. 1404–1419, 2019, doi: 10.1080/10447318.2018.1533152.
[52] Nigel Bevan, “Classifying and selecting UX and usability measures,” vol. 11, no. June, pp. 13–18, 2015, [Online]. Available: https://www.researchgate.net/publication/242077229_Classifying_and_selecting_UX_and_usability_measures.
[53] A. Seffah and E. Metzker, “The obstacles and myths of usability and software engineering,” *Commun. ACM*, vol. 47, no. 12, pp. 71–76, 2004, doi: 10.1145/1035134.1035136.
[54] R. E. Al-Qutaish and L. Cheikhi, “Analysis of the meanings of the measurement, assessment, and evaluation in iso/ieee software engineering standards,” *10th WSEAS Int. Conf. Appl. Comput. Appl. Comput. Sci. ACACOS’11*, pp. 110–115, 2011.
[55] C. Jooste, J. Van Biljon, and J. Mentz, “Usability evaluation guidelines for Business Intelligence applications,” *ACM Int. Conf. Proceeding Ser.*, pp. 331–340, 2013, doi: 10.1145/2513456.2513478.
[56] S. C. Ag, “ISO / IEC JTC1 / SC7 3932 Title Combined WD circulation and PDTR registration Combined WD circulation and PDTR registration , WD25060 , Software product Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for Usability — Genera,” *Syst. Eng.*, no. Resolution 937, 2008.
[57] P. Dey, M. Selvaraj, and B. Lee, *Robust user context analysis for multimodal interfaces*. 2011.
[58] P. Kortum and S. C. Peres, “The Relationship Between System Effectiveness and Subjective Usability Scores Using the System Usability Scale,” *Int. J. Hum. Comput. Interact.*, vol. 30, no. 7, pp. 575–584, 2014, doi: 10.1080/10447318.2014.904177.
[59] M. C. Schall, L. Cullen, P. Pennathur, H. Chen, K. Burrell, and G. Matthews, “Usability evaluation and implementation of a health information technology dashboard of evidence-based quality indicators,” *CIN - Comput. Informatics Nurs.*, vol. 35, no. 6, pp. 281–288, 2017, doi: 10.1097/CIN.0000000000000325.
[60] R. Carrillo, C. Renaud, Y. Prie, and E. Lavoue, “Dashboard for Monitoring Student Engagement in Mind Mapping Activities,” *Proc. - IEEE 17th Int. Conf. Adv. Learn. Technol. ICALT 2017*, pp. 433–437, 2017, doi: 10.1109/ICALT.2017.100.
[61] M. Gao, P. Kortum, and F. Oswald, “Psychometric evaluation of the USE (usefulness, satisfaction, and ease of use) questionnaire for reliability and validity,” *Proc. Hum. Factors Ergon. Soc.*, vol. 3, pp. 1414–1418, 2018, doi: 10.1177/1541931218621322.
[62] M. Kuhnel, L. Seiler, A. Honal, and D. Ifenthaler, “Mobile learning analytics in higher education: usability testing and evaluation of an app prototype,” *Interact. Technol. Smart Educ.*, vol. 15, no. 4, pp. 332–347, 2018, doi: 10.1108/ITSE-04-2018-0024.
[63] D. T. Y. Wu *et al.*, “Usability Testing of an Interactive Dashboard for Surgical Quality

Improvement in a Large Congenital Heart Center," *Appl. Clin. Inform.*, vol. 10, no. 5, pp. 859–869, 2019, doi: 10.1055/s-0039-1698466.
[64] M. E. Losa-Iglesias *et al.*, "The Usability of a Heartbeat Measuring Mobile Phone app: an Observational Study," *J. Med. Syst.*, vol. 43, no. 7, 2019, doi: 10.1007/s10916-019-1367-5.
[65] D. Dowding, J. A. Merrill, Y. Barrón, N. Onorato, K. Jonas, and D. Russell, "Usability Evaluation of a Dashboard for Home Care Nurses," *CIN - Comput. Informatics Nurs.*, vol. 37, no. 1, pp. 11–19, 2019, doi: 10.1097/CIN.0000000000000484.
[66] E. Chang and T. S. Dillon, "Automated Usability Testing," *Human-Computer Interact. INTERACT '97*, vol. 76, no. January, pp. 77–84, 1997, doi: 10.1007/978-0-387-35175-9_13.
[67] TechSmith, "Usability Testing with Morae," p. 57, 2009, [Online]. Available: https://assets.techsmith.com/Docs/pdf-morae/SW-Usability-Testing-with-Morae.pdf.
[68] N. Bevan and J. Carter, "Human-Computer Interaction. Theory, Design, Development and Practice," *Int. Conf. Human-Computer Interact. (pp. 268-278). Springer, Cham.*, vol. 9731, no. July, pp. 268–278, 2016, doi: 10.1007/978-3-319-39510-4.
[69] T. O.Nyumba, K. Wilson, C. J. Derrick, and N. Mukherjee, "The use of focus group discussion methodology: Insights from two decades of application in conservation," *Methods Ecol. Evol.*, vol. 9, no. 1, pp. 20–32, 2018, doi: 10.1111/2041-210X.12860.
[70] J. R. Lewis, "IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use," *Int. J. Hum. Comput. Interact.*, vol. 7, no. 1, pp. 57–78, 1995, doi: 10.1080/10447319509526110.
[71] J. P. Chin, V. A. Diehl, and K. L. Norman, "Development of an Instrument Measuring User Satisfaction of Human-Computer Interface," pp. 213–218.
[72] J. Kirakowski and M. Corbett, "SUMI: the Software Usability Measurement Inventory," *Br. J. Educ. Technol.*, vol. 24, no. 3, pp. 210–212, 1993, doi: 10.1111/j.1467-8535.1993.tb00076.x.
[73] J. R. Lewis, B. S. Utesch, and D. E. Maher, "UMUX-LITE - When there's no time for the SUS," *Conf. Hum. Factors Comput. Syst. - Proc.*, no. April 2013, pp. 2099–2102, 2013, doi: 10.1145/2470654.2481287.
[74] K. A. Ericsson and H. A. Simon, *Protocol analysis: Verbal reports as data.* Cambridge, MA, US: The MIT Press, 1984.
[75] D. Seminar and D. U. Experience, "User Experience White Paper," *Seminar*, pp. 1–12, 2011, [Online]. Available: http://www.allaboutux.org/files/UX-WhitePaper.pdf.
[76] P. Jenny, R. Yvonne, and S. Helen, *INTERACTION DESIGN beyond human–computer interaction*. 2015.
[77] M. Nosheen, Z. Sayed, M. S. Malik, and M. A. Fahiem, "An Evaluation Model for Measuring the Usability of Mobile Office Applications through User Interface Design Metrics," *Mehran Univ. Res. J. Eng. Technol.*, vol. 38, no. 3, pp. 641–654, 2019, doi: 10.22581/muet1982.1903.10.
[78] M. Georgsson and N. Staggers, "Quantifying usability: An evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics," *J. Am. Med. Informatics Assoc.*, vol. 23, no. 1, pp. 5–11, 2016, doi: 10.1093/jamia/ocv099.
[79] M. Oliveira, E. Cardoso, and M. Santana, "Measuring usability and dashboards efficiency: Framework developed using the principles of design of business intelligence and human machine interaction [Medição de usabilidade e eficiência de dashboards: Framework desenvolvida usando os princípios de des," *Atas da Conf. da Assoc. Port. Sist. Inf.*, vol. 15, no. Capsi 2015, pp. 338–362, 2015, doi: 10.18803/capsi.v15.338-362.
[80] D. Dowding, J. A. Merrill, Y. Barrón, N. Onorato, K. Jonas, and D. Russell, "Usability Evaluation of a Dashboard for Home Care Nurses," *CIN - Comput. Informatics Nurs.*, vol. 37, no. 1, pp. 11–19, 2019, doi: 10.1097/CIN.0000000000000484.
[81] J. Zhang and M. F. Walji, "TURF: Toward a unified framework of EHR usability," *J. Biomed. Inform.*, vol. 44, no. 6, pp. 1056–1067, 2011, doi: 10.1016/j.jbi.2011.08.005.
[82] O. M. Yigitbasioglu and O. Velcu, "A review of dashboards in performance management: Implications for design and research," *Int. J. Account. Inf. Syst.*, vol. 13, no. 1, pp. 41–59, 2012, doi: 10.1016/j.accinf.2011.08.002.
[83] O. Pastushenko, J. Hynek, and T. Hruška, "Evaluation of user interface design metrics using

generator of realistic-looking dashboard samples," *Expert Syst.*, no. September 2018, pp. 1–16, 2019, doi: 10.1111/exsy.12434.
[84] S. Few, "Tapping into the Power of Visual Perception," *Inf. dashboard Des. displaying data at-a-glance Monit.*, 2013, [Online]. Available: https://books.google.de/books/about/Information_Dashboard_Design.html?id=7k0EnAEACAAJ&redir_esc=y.
[85] E. Foster and R. Siddle, "The effectiveness of learning analytics for identifying at-risk students in higher education," *Assess. Eval. High. Educ.*, vol. 2938, 2019, doi: 10.1080/02602938.2019.1682118.
[86] G. Donahue, S. Weinschenk, J. Nowicki, C. De Araujo, D. D. C. Fettermann, and R. B. Bagno, "Usability is good business," *... //Www. Yucentrik. Ca/Usability. ...*, no. August, 1999.
[87] L. D. Roberts, J. A. Howell, and K. Seaman, "Give Me a Customizable Dashboard: Personalized Learning Analytics Dashboards in Higher Education," *Technol. Knowl. Learn.*, vol. 22, no. 3, pp. 317–333, 2017, doi: 10.1007/s10758-017-9316-1.
[88] T. Broos, K. Verbert, G. Langie, C. Van Soom, and T. De Laet, "Small data as a conversation starter for learning analytics: Exam results dashboard for first-year students in higher education," *J. Res. Innov. Teach. Learn.*, vol. 10, no. 2, pp. 94–106, 2017, doi: 10.1108/JRIT-05-2017-0010.
[89] E. Cardoso, T. Pedroso, F. Rações, A. Baptista, and J. Barateiro, "Learning Scorecard Dashboards : visualizing student learning experience," 2018.

Annex A

Table 25 summarizes the dimensions that were gathered from the literature review but discarded from the DATUS model and the reasons for it.

Table 25: Dimensions to measure usability discarded from the model

Dimension	Acceptance	Reason to discard
Accessibility	✓	-
Affect	×	To be measured in *Efficiency*
Appropriateness recognizability	✓	-
Attractive	×	To be measured in *User Interface Aesthetics*
Cognitive Workload	×	The dashboard allows decision making, but the decision is not made at the interface, it has an impact on the organisation/entity/user
Conformity	×	NA(*)
Context Coverage	×	To be measured in *Effectiveness*, *Efficiency* and *Satisfaction*
Control	×	To be measured in *Efficiency*
Effectiveness	✓	-
Efficiency	✓	-
Errors	×	To be measured in *Effectiveness*
Flexibility	×	To be measured in *Efficiency* and *Effectiveness*
Freedom from risk	×	NA
Global	×	To be measured in *Efficiency*
Helpfulness	×	To be measured in *Efficiency*, the level of available help
Interruptibility	×	NA
Learnability	✓	-
Memorability	×	NA
Operability	✓	-
Perceived usefulness	×	To be measured in *Satisfaction*, from the point of view that the user realises that it is useful
Productivity	×	NA
Safety	×	NA
Satisfaction	✓	-
Simplicity	×	To be measured in *Effectiveness*
Trustfulness	×	NA
Understandability	×	To be measured in *Learnability*
Universality	×	According to the author's definition, it is related and measured already in *Effectiveness*
Usable	×	To be measured in *Learnability* and *Efficiency*
Usefulness	×	To be measured in *Satisfaction*, from the point of view that the user realises that it is useful
User error protection	×	To be measured in *Effectiveness* and *Learnability*
User interface aesthetics	✓	-
Visibility	×	NA

(*) NA = Not Applicable

Annex B

DATUS questionnaire derived from two well-known post-study questionnaires: Computer System Usability Questionnaire (see **Figure 29**) and Questionnaire for User Interface Satisfaction (see **Figure 30**).

	The Computer System Usability Questionnaire Version 3		Strongly Agree						Strongly Disagree		NA
			1	2	3	4	5	6	7		
1	Overall, I am satisfied with how easy it is to use this system.		O	O	O	O	O	O	O		O
2	It is simple to use this system.		O	O	O	O	O	O	O		O
3	I am able to complete my work quickly using this system.		O	O	O	O	O	O	O		O
4	I feel comfortable using this system.		O	O	O	O	O	O	O		O
5	It was easy to learn to use this system		O	O	O	O	O	O	O		O
6	I believe I became productive quickly using this system.		O	O	O	O	O	O	O		O
7	The system gives error messages that clearly tell me how to fix problems.		O	O	O	O	O	O	O		O
8	Whenever I make a mistake using the system, I recover easily and quickly.		O	O	O	O	O	O	O		O
9	The information (such as online help, on-screen messages and other documentation) provided with this system is clear.		O	O	O	O	O	O	O		O
10	It is easy to find the information I needed.		O	O	O	O	O	O	O		O
11	The information provided with the system is effective in helping me complete my work.		O	O	O	O	O	O	O		O
12	The organization of information on the system screens is clear.		O	O	O	O	O	O	O		O
13	The interface* of this system is pleasant.		O	O	O	O	O	O	O		O
14	I like using the interface of this system.		O	O	O	O	O	O	O		O
15	This system has all the functions and capabilities I expect it to have.		O	O	O	O	O	O	O		O
16	Overall, I am satisfied with this system.		O	O	O	O	O	O	O		O

** The "interface" includes those items that you use to interact with the system. For example, some components of the interface are the keyboard, the mouse, the microphone, and the screens (including their graphics and language).*

Figure 29: The CSUQ (Version 3) [51].

OVERALL REACTION TO THE SOFTWARE		0	1	2	3	4	5	6	7	8	9		NA
1.	terrible	○	○	○	○	○	○	○	○	○	○	wonderful	○
2.	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
3.	frustrating	○	○	○	○	○	○	○	○	○	○	satisfying	○
4.	inadequate power	○	○	○	○	○	○	○	○	○	○	adequate power	○
5.	dull	○	○	○	○	○	○	○	○	○	○	stimulating	○
6.	rigid	○	○	○	○	○	○	○	○	○	○	flexible	○
SCREEN		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**
7. Reading characters on the screen	hard	○	○	○	○	○	○	○	○	○	○	easy	○
8. Highlighting simplifies task	not at all	○	○	○	○	○	○	○	○	○	○	very much	○
9. Organization of information	confusing	○	○	○	○	○	○	○	○	○	○	very clear	○
10. Sequence of screens	confusing	○	○	○	○	○	○	○	○	○	○	very clear	○
TERMINOLOGY AND SYSTEM INFORMATION		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**
11. Use of terms throughout system	inconsistent	○	○	○	○	○	○	○	○	○	○	consistent	○
12. Terminology related to task	never	○	○	○	○	○	○	○	○	○	○	always	○
13. Position of messages on screen	inconsistent	○	○	○	○	○	○	○	○	○	○	consistent	○
14. Prompts for input	confusing	○	○	○	○	○	○	○	○	○	○	clear	○
15. Computer informs about its progress	never	○	○	○	○	○	○	○	○	○	○	always	○
16. Error messages	unhelpful	○	○	○	○	○	○	○	○	○	○	helpful	○
LEARNING		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**
17. Learning to operate the system	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
18. Exploring new features by trial and error	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
19. Remembering names and use of commands	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
20. Performing tasks is straightforward	never	○	○	○	○	○	○	○	○	○	○	always	○
21. Help messages on the screen	unhelpful	○	○	○	○	○	○	○	○	○	○	helpful	○
22. Supplemental reference materials	confusing	○	○	○	○	○	○	○	○	○	○	clear	○
SYSTEM CAPABILITIES		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**
23. System speed	too slow	○	○	○	○	○	○	○	○	○	○	fast enough	○
24. System reliability	unreliable	○	○	○	○	○	○	○	○	○	○	reliable	○
25. System tends to be	noisy	○	○	○	○	○	○	○	○	○	○	quiet	○
26. Correcting your mistakes	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
27. Designed for all levels of users	never	○	○	○	○	○	○	○	○	○	○	always	○
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**

List the most **negative** aspect(s):

1. _____
2. _____
3. _____

List the most **positive** aspect(s):

1. _____
2. _____
3. _____

Figure 30: The QUIS

(Source: https://garyperlman.com/quest/quest.cgi?form=QUIS)

Annex C

Table 26 summarizes features from student dashboards reviewed in the literature review.

Table 26: Summary of **s**tudent dashboard features from the literature review

Authors	Participants		Features	Adapted ✓/ ×	Exclusion criteria and additional notes
	#	Type			
[87]	199	Students	A reminder of when upcoming assignments are due	✓	
			Links to additional resources	✓	
			Student's grades	✓	
			Program guidance (i.e., suggested readings)	×	Given with the class's materials
			Early alerts about performance	×	Considered as information on grades
			Support services and their contact details	×	Present on the website of Iscte and Fénix platform
			A calendar	✓	
			Student's attendance	✓	
			A link to provide feedback to teaching staff	×	Given as pedagogical surveys every semester
			A colour system indicating if the student is doing well or not	✓	Not specifically as a widget but along the dashboard
			Student's grades in comparison to peers	✓	
			Student's class participation	×	Impracticable
[39]	10	Teachers	Number of times student access the course forum	✓	Adapted to the e-learning and Fénix platforms
			Amount of forum contributions	×	Not applicable
			Number of times students consult the course student guide	×	Not applicable
			Number of times each student has accessed the course in one week	✓	Adapted to the e-learning and Fénix platforms
			Number of activities that have been delivered out of date	✓	
[85]	+7000	Staff and Student	Contextual data about the student (photo, course details)	✓	
			Written notes by university staff members, for example, tutors	×	Not applicable
			Early warning alerts to the students' tutor after 14 days of inactivity	×	Not applicable

Authors	Participants		Features	Adapted ✓/ ×	Exclusion criteria and additional notes
	#	Type			
			Attendance monitoring	✓	
			Card accesses to buildings	×	Impracticable
			E-resource logins (e.g. online books and journals)	×	Impracticable
			Library loans	✓	Adapted to status of requisitioned books
			Virtual learning environment logins	✓	
			Virtual learning environment learning room use	×	Impracticable
			Online coursework (assessments) submission	✓	Adapted as reminders of assessments submissions as milestones
[88]	887	Students	Course grades and its comparison with peers	✓	
			Overall achievement (given by their peers)	✓	
			Tips and tricks (to help improve student's learning study behaviour)	✓	
			Regulations (provides a highlight of the regulations that matter)	×	Related with indications for the exams, given by the professors or by the e-learning platform
[38]	7028	Students	Grades and students position compared with peers	✓	
			Course result	✓	
			Credits achieved in percentage	✓	
			Failed courses student's position against peers	✓	Adapted
			Length of bachelor in years (prediction with historical data)	✓	
[89]	112	Students	Personal information	✓	
			Collected badges versus the total amount	×	Adapted to progress in weeks or attendance in classes
			Student performance in comparison with the average performance of the class	✓	
			Quests grades in XP (gamification)	×	Not applicable

Annex D

Table 27 presents the list of tasks in Portuguese, i.e. the original version given to the participants.

Table 27: List of tasks (original version in Portuguese)

Task-id	Task description
1	Posso alterar o ano letivo da informação que estou a visualizar no dashboard?
2	A minha licenciatura é…
3	A minha média é de…
4	Estou inscrito no Iscte desde…
5	O ano que frequento é o…
6	Quantos créditos já fiz este ano letivo?
7	Relativamente ao estado das propinas, estão em dia?
8	Quantas unidades curriculares aguardam o lançamento da nota?
9	A quantas unidades curriculares reprovei?
10	A quantas unidades curriculares fui aprovado?
11	A minha nota mais alta foi na unidade curricular de...
12	No primeiro semestre, quantas cadeiras tenho em atraso?
13	Consigo alterar o tipo de gráfico do desempenho comparado com os meus colegas?
14	A média mais frequente entre os meus colegas de curso é de...
15	Em média, os alunos licenciados em IGE terminam o curso em quanto tempo?
16	Quantos livros requisitaste neste ano letivo?
17	O número médio de acessos ao Fénix dos teus colegas de curso na semana 5 foi de…
18	Consigo consultar a informação detalhada sobre o segundo semestre?
19	O dia em que estou a consultar o dashboard é…
20	Em quantas unidades curriculares do ano passado estás incrito neste ano letivo?
21	A próxima milestone pertence à unidade curricular de…
22	A unidade curricular em que já terminei todas as milestones é…
23	A última milestone de Macroeconomia é…
24	A média de Excel Avançado no ano de 2017/2018 foi de…
25	Consigo consultar a média de qualquer unidade curricular desde o ano letivo de…
26	Na UC de Conceção e Desenvolvimento de Sistemas de Informação, qual foi a taxa de aprovação real no ano letivo de 2017/2018?
27	Em quantas unidades curriculares tenho mais de 7 dias até à próxima milestone?
28	Em quantas unidades curriculares estou em risco de reprovar por causa da assiduidade?
29	No ano letivo anterior, ou seja 2018/2019, a taxa de aprovação real na UC de Complementos de Contabilidade Financeira foi de…
30	Consigo encontrar onde posso ter mais informação sobre a taxa de aprovação total?
31	Quantas semanas faltam para terminar o semestre?
32	A quantos inquéritos já respondi?
33	Tenho algum livro requisitado cuja data limite de entrega já tenha passado?
34	Consigo encontrar as dicas de antigos alunos?

After completing the task list, the participants answered the DATUS questionnaire and two open-ended questions whose answers were aggregated in **Table 28** and **Table 29**.

Table 28: Participants' suggestions of new features for the dashboard

Is there any information you would like to see that has not been considered on the dashboard?
No espaço das assiduidades, não percebi se é a "minha assiduidade VS media meus colegas" ou "minha assiduidade VS minimo", mas seria útil ter noção dos mínimos Seria interessante ter uma lista dos livros requesitados pelos meus colegas de curso, do mais requesitado ao menos, e tambem se os livros em questão estao disponiveis
Horário
Adicionaria a opção de colocar eventos no calendário, com atividades extra curriculares da faculdade (exemplo: jogo de vólei,…).
Penso que as informações no dashboard são relevantes, não encontro mais informações que pudessem ser incluídas.
Logo no início, havendo propinas em atraso, poderiam aparecer logo a referência de pagamento
Na parte da assiduidade ter a informação também expressa em horas.
Data Milestones
Link direto para o pagamento das propinas
"Poderia ter um widget com informações importantes da faculdade, como por exemplo um "Hoje, no Iscte" ou algo que mostrasse avisos importantes como "Durantes os próximos 3 dias, o Refeitório vai fechar mais cedo", etc.
Poderia também ter outro widget com informação das atividades extracurriculares que existem no Iscte. Algo como um widget onde fossem passando várias imagens (e.g., AE, Tuna, Desporto universitário, etc) e quem quuisesse saber mais carregava na foto e encaminhava para os respetivos sites ou páginas."
Não, apenas tiraria a informação acerca da atividade no Fénix
Acredito estarem todas as informações essenciais para o aluno. Apenas considero que a secção da "Actividade Online", relativamente à actividade do Fénix, talvez não seja tão relevante na ótica de um aluno e sim mais na de um professor.
Em cada UC ter um link para a FUC da mesma.
No calendário das milestones para além de dizer que foi terminada podia ter a nota com que foi concluída.

Table 29: Participants' comments and suggestions on the dashboard proposal

Do you have any other comments or suggestions you want to share with us?
Não achei muito interessante as dicas para os alunos
A atividade online do Fénix, pessoalmente não seria algo que iria usar mas também não me incomoda.
A ideia de ter noçao da minha media em relaçao à dos meus colegas é bastante boa Principalmente as medias, mas tambem as entradas no Fenix, e requesitos de livros Gosto bastante da facilidade em visualizar as datas dos milestones No geral, toda a informaçao do ""resumo do ano letivo"" é bastante util Colocava a informaçao dos livros requesitados, numa pagina mais geral (por ex. "ano letivo")
"Atividade Online de pouca utilidade Mudaria o aspeto dos milestones"
Fico um pouco confusa com os gráficos média ponderada e distribuição da média ponderada = não percebo bem a diferença
Poderiam existir mais separadores para a informações estar exposta com mais espaço.
Muito acessível, senti que podia ser uma excelente ferramenta para o Fenix, pois tem as informações importantes contidas de forma organizada e visualmente apelativa.

Talvez a secção "atividade no fénix" não seja muito útil.
A atividade online e o nr de livros requisitados pelos colegas talvez não seja muito relevante
Mostrar as horas a que a informação é apresentada (ou seja: Última atualização: 12:32:00).
Clarificar qual o semestre em que estamos; Link para mais informações sobre a UC; Notas em escala internacional; Reconhecimentos / prémios de mérito.
Alguma da informação menos relevante como por exemplo o número de semanas que faltam até ao fim do semestre, encontrava-se no meio do ecrã, quando seria de esperar que estivesse nas laterais ou perto da data. Achei que os vários gráficos que expunham a nossa informação em relação à média são bastante úteis. Considero que as dicas de antigos alunos poderão não ser muito úteis em certas cadeiras devido às alterações ao longo dos anos.
Pessoalmente, a parte das taxas de aprovação e os gráficos de desempenho comparado com os colegas não me despertam grande interesse. Por outro lado, gostei bastante do calendário e da assiduidade e adorei o facto de termos a informação relativa aos empréstimos da biblioteca (especialmente a parte do estado e das datas limites de entrega, para mim isso seria mesmo muitíssimo útil).
Quanto ao número de matrículas, considero que pode funcionar como um fator motivador para o próprio aluno olhar esse número e comparar com o ano que frequenta, mas para alguns também pode ser desmotivador relembrar o seu insucesso, se assim for o caso.
No quadro das Milestones, ter a data quando é posto o rato por cima das Milestone futuras.
O calendário de Milestones, inicialmente, parece algo confuso. Quando foi pedida a task 27 ainda demorou a processar que tinha de encontrar um ponto cinzento depois de um azul, que isso significava que a próxima era passados 7 dias. Está bem feito e catchy em termos visuais, mas demora-se um pouco mais a compreender como funciona.
Foi na página do 2º semestre que senti maiores dificuldades iniciais dada a elevada quantidade de informação nova. Assim que percebi o que cada parâmetro significava pareceu-me incrível e tudo muito útil (especialmente as milestones)
Não tenho interesse em ter o quadro "Desempenho Comparado com os Colegas" e a informação "Tempo médio de conclusão de curso" constantemente apresentada na página principal, pois poderia ser uma informação que consultaríamos em situações oportunas.
Na secção "Biblioteca", o esquema de cores não segue o mesmo padrão que as restantes secções do dashboard. Isto é, ao longo do dashboard recorre-se ao uso da cor (azul) para dar destaque à informação relativa ao aluno, que quando comparada com dados relativos a colegas, é mais importante. No entanto, na secção da "Biblioteca", os dados destacados com a cor azul dizem respeito ao dos colegas e não aos do utilizador.
A vizualização do dia no calendário não foi intuitiva. Retiraria as dicas de antigos alunos só para facilitar a logística de processamento de informação
Na minha opinião não acho relevante as dicas de antigos alunos.
Os gráficos calendário e assiduidade são bastante úteis e tornam o nosso trabalho mais eficiente para além de darem uma visão geral de todos os trabalhos pendentes do semestre. O facto do separador ter todas as cadeiras "inscritas" do semestre independentemente do ano.
Colocar os anos por ordem crescente (separador ano letivo). Em suma, é muito útil esta forma de vizualização
Actividades extracurriculares, eventos do iscte , milestones personalizáveis Não considero pertinente a comparação de média com os colegas, no entanto caso isso houvesse seria mais interessante comparar a media com todos os anos anteriores e não apenas com os colegas atuais.

Annex E

The following **Table 30** identifies the literature review of metrics to measure the correspondent dimensions of usability evaluation.

Table 30: Metrics to measure dimensions from literature review

Dimension	Metrics
Effectiveness	Success or failure rate
	Tasks completed
	Ability to completely perform assigned tasks
	Task completion
	Objectives achieved
	Accuracy of tasks completion
	Performing necessary steps to complete given tasks
	Error rate
	Errors in a task
	Tasks with errors
	Number of errors
	Number of user errors
	Number of navigational errors
Efficiency	Time to complete the task (human resources), materials, or the financial cost of usage
	Time taken for the completion of tasks within given time
	Time to execute a task
	Time to complete task
	Time taken to find desired result(s)
	Task time
	Time efficiency
	Minimal action
	Minimal memory load
	Minimal number of actions
	Unnecessary actions
	Fatigue
	Navigability
	Navigation (move around easily)
	Help (the degree they can use it without need for assistance)
	Feedback
	Task success rate
	Number of tasks completed successfully
	Goal time
	Goal efficiency by correct answers
	Goal efficiency by incorrect answers
Satisfaction	Overall satisfaction
	Satisfaction with features
	Feature utilisation
	Proportion of user complaints about a particular feature
	User trusts
	User pleasure
	Physical comfort
	Usefulness
	Trust
	Pleasure
	Comfort
	Productivity
	Resource utilization

	Loading time
	Level of agreement
	Feel happy while interacting with the application interface
	Attitude of user towards the application
	User experience while performing given tasks
	Opinion and Suggestion
	Rating through survey/interview
Learnability	Task completeness and the usage of documentation and help resources
	Time to discover how to perform a task
	Task time of the first use
	Number of trials to solve task
	Minimal action
	Minimal memory load
	User guidance
	Self-explanatory user interface
	Self-descriptiveness
	Simplicity
	Familiarity
	Learning curve (number of navigational errors)
	Easy to learn icons and tabs
Accessibility	Task completeness and accessibility issues
	Users Comments
	Flexibility
	Minimal action
	Minimal memory load
	Operability
	User guidance
	Consistency
	Self-descriptiveness
	Readability
	Navigability
	Simplicity
Appropriate recognizability	Description completeness
	Demonstration coverage
	Entry point self-descriptiveness
Operability	Operational consistency
	Message clarity
	Functional customizability
	User interface customizability
	Monitoring capability
	Undo capability
	Understandable categorization of information
	Appearance consistency
	Input device support