ADEDU

ACCESSIBLE DIGITAL EDUCATION

# ADEDU

## Accessible Digital EDUcation - ADEDU

# D2.4 Evaluation report of ADA digital environment

Hellenic Open University

ENHANCING DIGITAL SKILLS ACROSS EUROPE

HELLENIC OPEN UNIVERSITY

EASPD

iasis

MEDIA EDUCATION CENTRE

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EACEA. Neither the European Union nor the granting authority can be held responsible for them. Proposal number: 101133970.

Co-funded by the European Union




Co-funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

# Table of Contents

# 1. Introduction

In today's digital age, Massive Open Online Courses (MOOCs) have revolutionised education by providing accessible, flexible, and cost-effective learning opportunities to a global audience. The ALL DIGITAL Academy (ADA) Platform, is a training hub where free online courses and open-access resources are provided. ADA plays a crucial role in this educational transformation. However, to maintain its competitive edge and ensure user satisfaction, it is imperative to conduct thorough usability, including accessibility, and user experience (UX) evaluations.

**Usability and UX evaluations are essential for several reasons**. Firstly, they help identify potential barriers that learners may encounter while navigating the platform. These barriers can range from complicated registration processes to difficult-to-find course materials. **By addressing these issues, the platform can enhance its overall usability, making it easier for users to access and benefit from the educational resources available**.

Secondly, a positive user experience is critical in retaining learners and encouraging course completion. MOOCs often face high dropout rates, which can be attributed to frustrating user interfaces and poor design. A comprehensive UX evaluation can uncover aspects of the platform that may detract from the learning experience, such as unintuitive navigation, slow load times, or lack of engaging content. Improving these aspects can significantly boost user engagement and satisfaction.

Moreover, usability and UX evaluations provide valuable insights into user behavior and preferences. Understanding how learners interact with the platform enables developers to tailor the experience to better meet user needs. This user-centered approach can lead to innovative features and improvements that enhance the educational experience.

Finally, **in a competitive market, the usability and UX of a MOOC platform can be a differentiator**. Platforms that offer seamless, enjoyable experiences are more likely to attract and retain users. Regular evaluations ensure that the ADA Platform remains user-friendly and responsive to changing user expectations and technological advancements.

In conclusion, conducting usability and UX evaluations for the ADA Platform is vital for ensuring its effectiveness, user satisfaction, and continued success in the dynamic field of online education.

## 1.1. Usability

According to the ISO 9241-11 usability is defined as "*the extent to which a product can be used by specific users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use*".

In specific, effectiveness shows the accuracy with which users achieve certain goals and is typically measured using error rates and task success rates (Albert and Tullis 2013; Hornbæk and Law 2007). Efficiency is related to resources spent by users in order to complete a task and is typically measured through time on task and learning time. Both effectiveness and efficiency measurements represent the objective aspect of usability.

Usability evaluation studies are mainly based on two categories of data: (a) task performance (e.g. task completion time and success rate) and (b) self-reported satisfaction, which are typically collected through a questionnaire.

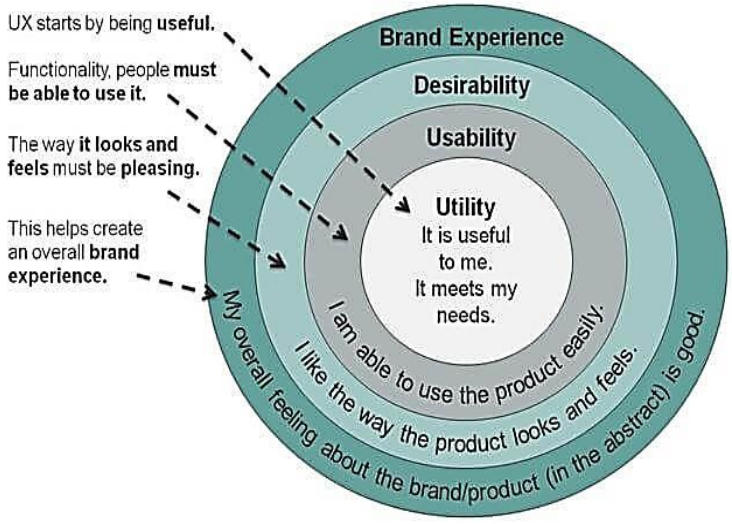



*Figure 1. The layered approach highlights how fundamental aspects of product design contribute to the broader perception of the brand. Source: User Experience 2008 conference by the Nielsen Norman Group in Amsterdam*

User satisfaction is connected to their overall attitude to the product and is typically captured using psychometric scales (Hornbæk and Law 2007). System Usability Scale (SUS) (Brooke 1996), User Engagement Scale (UES) (O'Brien and

Toms 2010) and Questionnaire for User Interface Satisfaction (QUIS) (Chin, Diehl, and Norman 1988) are also proposed to measure satisfaction, to some extent, however they still cannot fully capture fuzzy and dynamic aspects of interaction experience such as aesthetics, emotions, hedonic, experiential etc.

## 1.2. User eXperience (UX)

User eXperience (UX) is concerned with the entire process of acquiring and integrating a product, including aspects of branding, design, usability and function. UX begins before the product is even in the user's hands.

In specific, UX aims at the prevention of frustration and dissatisfaction (Hassenzahl and Tractinsky 2006). Hence, developing for UX requires a deep understanding of how users feel during their interaction with a system or a product. **As a recently established research field there has not been a widely accepted definition for the UX term**. Some of the most dominant perspectives about UX are the following:

- Bevan (2009) describes UX as "an elaboration of the satisfaction component of usability".
- According to ISO 9241 - 2010, UX is defined as "a person's perceptions and responses that results from the use or anticipated use of a product, system or service".
- Wechsung et al. (2011) describes UX using two aspects: a) "ease of use", and b) "joy of use" which both can be used to determine user satisfaction.
- Norman & Nielsen (2014) posit that "UX encompasses all aspects of the end-user's interaction with the company, its services, and its products".

## 1.3. MOOC Platforms Evaluation

Evaluation of a MOOC platform could reveal issues that should be addressed either on the interface or in the content. A user-friendly and intuitive MOOC platform could positively contribute to learners' engagement during a course delivery (Vercellotti 2018). On the contrary, poor design and usability issues of user interfaces can negatively affect e-learning (Zaharias and Poylymenakou 2009).

Existing research typically evaluates the usability among popular platforms dedicated to developing MOOCs. Tsironis, Katsanos, and Xenos (2016) investigated the usability of three MOOC platforms (i.e. edX, Coursera, and Udacity). Results revealed that Coursera was significantly more usable compared to the other two (edX and Udacity). Regarding task-related metrics, results showed that users' mean task time was significantly affected by the platform. Electronic learning platforms are evolving and their evaluation is becoming more complex and challenging with time. Yet, the evaluation of electronic learning services is intrinsically linked to improving the performance of documentation services.

Similarly, a study focused on the instructional and interface design aspects of the MOOC platform Udacity and evaluated its usability by using a questionnaire and task scenario (Anyatasia, Santoso, and Junus 2020). Korableva et al. (2019) investigated the usability of MOOCs offered by Coursera and Open Education, which is the national online educational platform of Russia. The participants assessed the interfaces through a questionnaire and adjective description. Results showed higher usability for Coursera than Open Education. Suggestions for interface improvements were also reported for both platforms.

There is also interest in evaluating the UX in localised MOOCs, as it has been done for two MOOCs on the Chinese University MOOC platform and Coursera (Liu et al., 2020). It was found that the UX in the MOOCs offered by the Chinese

University MOOC platform might also depend on culture-preferred interface design.

# 2. Methodology - Procedures

## 2.1. Participants

A sample of 13 (8 males, 4 females, 1 did not answer) users was recruited to participate in the usability study. Regarding participants' age the most of them belong to 34 - 44 group. Out of these: 9 are educators (as per the *Accessible Digital Education* project end users) of different educational levels, while 5 of them are also facing learning and other disabilities (specifically, two of them learning disabilities, i.e. Attention Deficit Hyperactivity Disorder (ADHD), and other multiple disabilities and two others visual impairment issues). Figure 2 illustrates some basic demographic information about the participants.

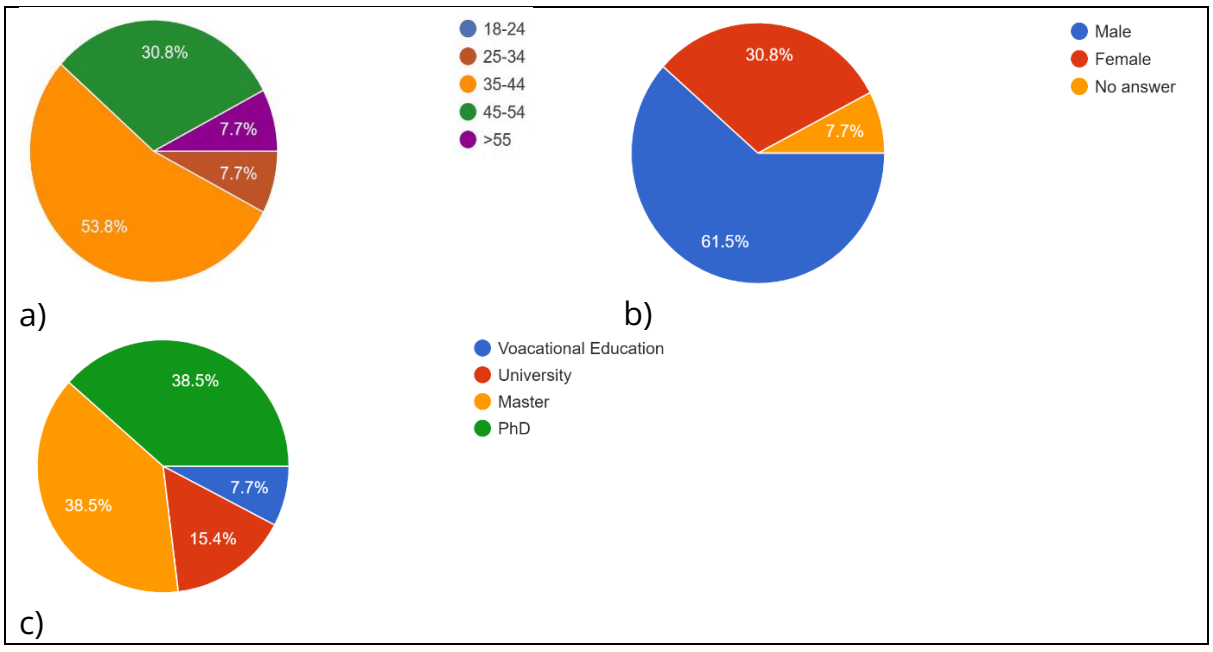



*Figure 2. Demographic info: a) Age, b) gender, and c) Educational level*

## 2.2. Interaction Tasks

In the present study, participants were asked to perform a set of typical interaction tasks within ADA platform. Some of the selected tasks were also used in (Liapis et al. 2023; Tsironis et al. 2016). Tasks were designed to require adequate navigation.

More specifically, participants were asked to:

- Navigate in the platform in order to Send a private message to **Alexandros Liapis**. **Subject**: Hi – **Message**: Hello Alexandros. (Task 1).
- Navigate in the platform in order to find and read the research paper entitled "**Artificial Intelligence for HCI: A Modern Approach**" (Task 2).
- Navigate in the platform in order to find what was the starting date of "**The GenAIEdu MOOC at a glance**" training program (Task 3).
- Navigate in the platform in order to find the webinar "***AI for problem solving & the Certification Capstone***". What is the duration of the video? (Task 4).
- Navigate in the platform in order to find the learning objectives of ***AI MOOC*** (Task 5)

## 2.3. Experimental Sessions

Each participant was asked to fill out an appropriate consent form before participating in the study.

Next, important information were given. More specifically, participants were informed about:

1. **No Right or Wrong Answers**: Remember, we are testing the product, not you. There are no right or wrong answers or actions.
2. **Think Aloud**: Please share your thoughts as you navigate through the tasks. Verbalising your thought process helps us understand your decisions and reactions.
3. **Ask Questions**: If you have any questions or need clarification during the tasks, feel free to ask.
4. **Time Limit**: To keep the session on track, each task should be completed within 5 minutes. If you are unable to finish a task within this time frame, don’t worry—just let us know, and we'll move on to the next task.

Subsequently, the interaction scenarios were presented to participants in a counterbalance mode to reduce order effects.

At the end of interaction with each platform, participants answered the Greek version of the System Usability Scale (SUS) and the AttrakDiff. Each experimental session lasted approximately 40 minutes.

At the end of the session, each participant was informed about the purposes of the study.

## 3. Perceived Usability – System Usability Scale (SUS)

The System Usability Scale (SUS) is a widely-used tool in User eXperience (UX) research to assess the usability of a product, service, or system. Developed by Brooke (1996), SUS provides a quick and reliable way of measuring the perceived usability of interfaces through a 10-item questionnaire . Each item is scored on a 5-point Likert scale ranging from "Strongly Agree" to "Strongly Disagree."

One of the key strengths of SUS is its simplicity and flexibility. The 10 questions cover a broad range of usability aspects, such as ease of use, the need for support, and the consistency of the system. After respondents complete the questionnaire, their scores are converted to a 0-100 scale, where higher scores indicate better usability. A score above 68 is generally considered above average and constitutes a minimum acceptable threshold (Bangor, Kortum, and Miller 2008). In our study mean SUS score is **74,2** with SD=**21,1**. Such score indicates good to excellent usability performance.

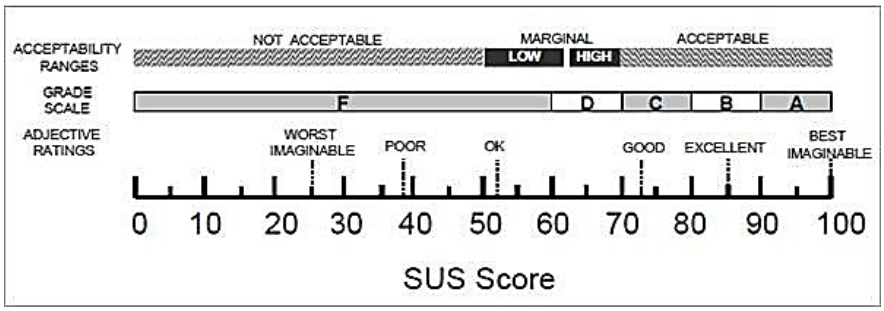



*Figure 3. SUS score levels (Bangor et al. 2008).*

The reliability and validity of SUS have been extensively validated in various studies. Research has shown that SUS can effectively differentiate between usable and unusable systems, making it a valuable tool for both formative and summative evaluations (Sauro 2011a). Its adaptability allows it to be applied across different types of systems, from software applications and websites to hardware and consumer products (Lewis and Sauro 2009). Moreover, SUS has the advantage of being relatively easy and quick to administer, which makes it

suitable for both small and large sample sizes. This efficiency does not compromise its accuracy, as SUS has been proven to produce consistent and reliable results. Its wide acceptance in the industry further reinforces its credibility as a standard usability evaluation method (Stetson and Tullis 2004).

Despite its strengths, it is essential to recognise that SUS provides a high-level overview of usability and may not identify specific issues. Therefore, it is often used in conjunction with other usability testing methods to obtain a comprehensive understanding of user experience (Sauro 2011a).

## 3.1. Perceived Usability and Learning Platforms

The System Usability Scale (SUS) has been widely utilised in various studies to assess the usability of learning platforms (Liapis et al. 2023; Tsironis et al. 2016; Vlachogianni and Tselios 2021). Using SUS in learning platforms provides a structured approach to understanding and improving usability. By regularly applying SUS, developers and educators can ensure that the platform evolves in a way that best supports the learning experience.

## 3.2. Data Analysis

Analyzing responses from a SUS questionnaire involves several steps to ensure the results are interpreted correctly and provide actionable insights. Here is a commonly used approach to analysing SUS questionnaire responses:

**1. Data Collection and Preparation**

- **Gather Responses**: Collect completed SUS questionnaires from participants. Ensure that each questionnaire is fully completed to maintain the integrity of the data. In order to meet the specific requirements, Google Forms have been used.

**2. Scoring the SUS**

Each SUS questionnaire consists of 10 items, with responses on a 5-point Likert scale ranging from "Strongly Disagree" (1) to "Strongly Agree" (5). The scoring process involves the following steps:

- **Odd-Numbered Questions**: Subtract 1 from the score.
- **Even-Numbered Questions**: Subtract the score for each even-numbered question from 5.
- **Summing Scores**: Sum the adjusted scores for all 10 items.
- **Scaling**: Multiply the total score by 2.5 to convert it to a scale of 0 to 100.

**3. Interpretation of SUS Scores**

- **Overall Usability**: SUS scores range from 0 to 100. A score above 68 is generally considered above average, indicating good usability, while scores below 68 suggest areas for improvement.

## 3.3. SUS Score - Results

Participants navigated in the ADA platform and performed five tasks. Next, participants were asked to complete the SUS questionnaire. All in all, **130 responses were collected**. In Table 1 we present the responses distribution. In addition, Figure 4 depicts participants' responses to each SUS question, broken down by scale item (Strongly Disagree, Disagree, Neither Agree Nor Disagree, Agree, Strongly Agree).

The majority of responses fall into the "Agree" category (30.00%), **indicating a generally positive perception of the system's usability**. There are significant portions of "Strongly Disagree" (23.08%) and "Disagree" (14.62%) responses, **suggesting some dissatisfaction or difficulty with certain aspects of the system**.

*Table 1. Number of responses for each question per scale item.*

| | **Strongly Disagree (1)** | **Disagree (2)** | **Neither Agree Nor Disagree (3)** | **Agree (4)** | **Strongly Agree (5)** |
|---|---|---|---|---|---|
| *"I think that I would like to use this system frequently."* | 0 | 2 | 1 | 8 | 2 |
| *"I found the system unnecessarily complex."* | 3 | 5 | 3 | 2 | 0 |
| *"I thought the system was easy to use."* | 0 | 1 | 3 | 5 | 4 |
| *"I think that I would need the support of a technical person to be able to use this system."* | 7 | 4 | 1 | 1 | 0 |
| *"I found the various functions in this system were well integrated."* | 0 | 0 | 4 | 6 | 3 |
| *"I thought there was too much inconsistency in this system."* | 6 | 2 | 4 | 1 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| *"I would imagine that most people would learn to use this system very quickly."* | 0 | 0 | 2 | 9 | 2 |
| *"I found the system very cumbersome to use."* | 6 | 1 | 4 | 1 | 1 |
| *"I felt very confident using the system."* | 1 | 0 | 2 | 5 | 5 |
| *"I needed to learn a lot of things before I could get going with this system."* | 7 | 4 | 1 | 1 | 0 |
| *total* | 30 | 19 | 25 | 39 | 17 |
| *Total (%)* | 23,08% | 14,62% | 19,23% | 30,00% | 13,08% |

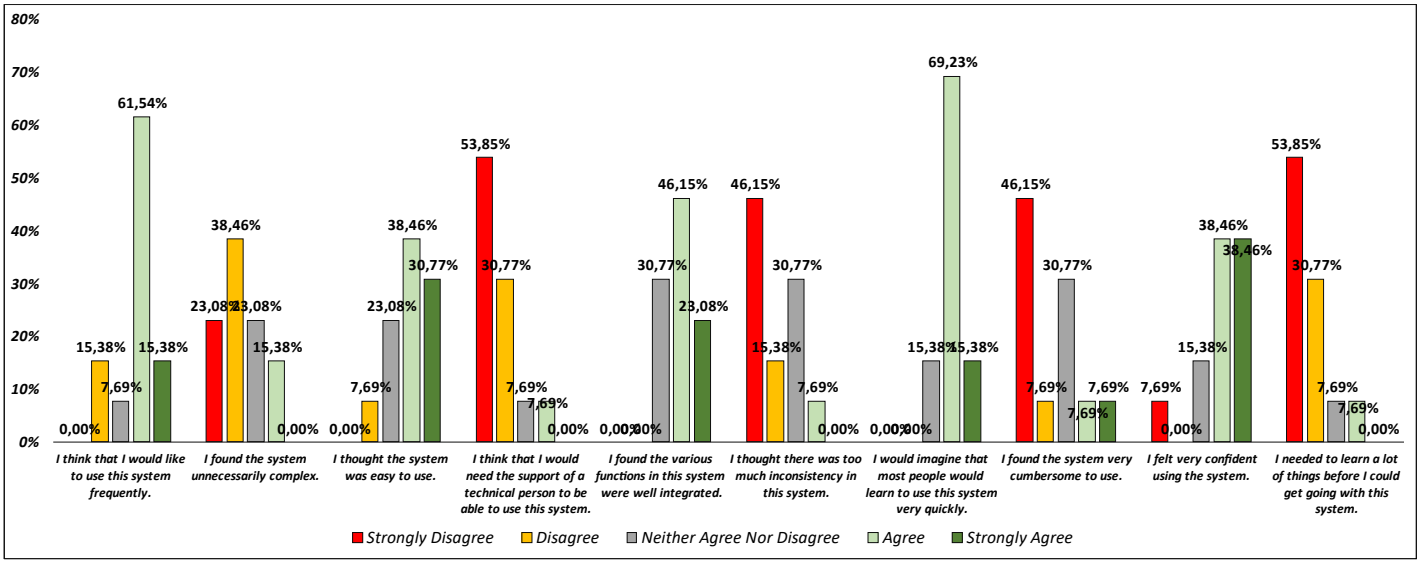



*Figure 4. A per cent representation of participants' responses to each SUS question, broken down by scale item (Strongly Disagree, Disagree, Neither Agree Nor Disagree, Agree, Strongly Agree).*

### 3.3.1. Positive Aspects

*Frequent Use*: Most participants (66.67% combining "Agree" and "Strongly Agree") are **inclined to use the system frequently**, indicating overall acceptance and potential for regular use.

*Ease of Use*: A notable 75% of participants (combining "Agree" and "Strongly Agree") **find the system easy to use, highlighting its user-friendly design**.

*Integration of Functions*: The majority (69.23%) agree that the system's functions are well integrated, which is crucial for seamless user experience.

### 3.3.2. Areas for Improvement

*Complexity and Consistency*: **There are significant concerns regarding complexity** (38.46% disagree) and **inconsistency** (61.54% disagree), suggesting that the system may need simplification and better consistency.

### 3.3.3. Confidence and Learning Curve

*Confidence*: There is a mixed response regarding user confidence, with 50% expressing confidence but some (23.08%) indicating a lack of it.

*Learning Requirements*: A significant number of participants (84.62%) feel they do not need to learn a lot before using the system.

Finally, Table 2 presents the mean score and standard deviation for each question of SUS score.

*Table 2. Mean score for each SUS question.*

| Question | Mean | SD |
|---|---|---|
| *I think that I would like to use this system frequently.* | 3,77 | 3,13 |
| *I found the system unnecessarily complex.* | 2,31 | 1,82 |
| *I thought the system was easy to use.* | 3,92 | 2,07 |
| *I think that I would need the support of a technical person to be able to use this system.* | 1,69 | 2,88 |
| *I found the various functions in this system were well integrated.* | 3,92 | 2,61 |
| *I thought there was too much inconsistency in this system.* | 2,00 | 2,41 |
| *I would imagine that most people would learn to use this system very quickly.* | 4,00 | 3,71 |
| *I found the system very cumbersome to use.* | 2,23 | 2,30 |
| *I felt very confident using the system.* | 4,00 | 2,30 |

| | | |
|---|---|---|
| *I needed to learn a lot of things before I could get going with this system.* | 1,69 | 2,88 |

## 3.4. Open-ended Questions

At the end of SUS questionnaire participants were able to answer in three open-ended questions. More specifically:

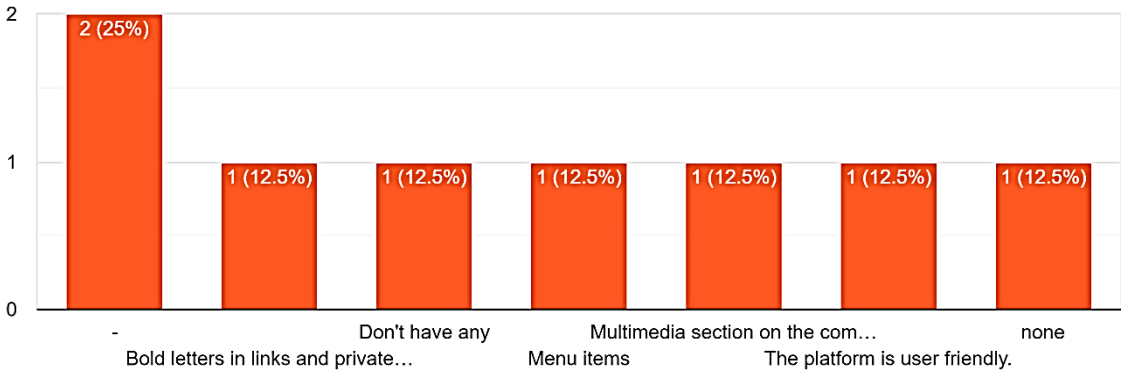
*What aspects of our platform do you find most effective and beneficial, and why do you think they work well? – (7 responses)*

- **The "search" function** and the tabs at homepage
- intuitive and easy-to-navigate interface
- The OER section because it provides a lot of extra material on the training subjects
- The very well integrated **search function**
- Everything was O.K.
- The **search button** was useful because I couldn't manually find what I was looking for.
- The **search button**
- **Search** works well

*What challenges or issues have you encountered with our platform, and how have they impacted your experience? - (6 responses)*

- it was an excellent experience
- Mostly with user based actions e.g. finding someone to send message as a minor issue. A small impact to the whole experience
- Not actual challenges were met
- I could not find the names of some participants in members' section
- None. Everything was very clear to me.
- It was confusing quickly finding the links to open the articles because they weren't highlighted with bold letters.

*What specific changes or improvements would you suggest to enhance our platform, and how do you believe these changes would make a difference?*



# 4. Evaluating Products' Attractiveness: The Attrakdiff approach

Evaluating User eXperience (UX) is crucial for the design and improvement of products and services. One of the most popular tools for UX evaluation is AttrakDiff, developed by Hassenzahl, Burmester, and Koller (2003). AttrakDiff is a questionnaire that uses bipolar adjective pairs to measure the perceived quality of interactive products.

AttrakDiff is based on two main dimensions: **pragmatic quality** and **hedonic quality** (Figure 5). More specifically, pragmatic quality assesses the usability and functionality of the product, while hedonic quality measures the pleasure and aesthetic experience the product offers. These two dimensions aim to provide an overall evaluation of the product's attractiveness.

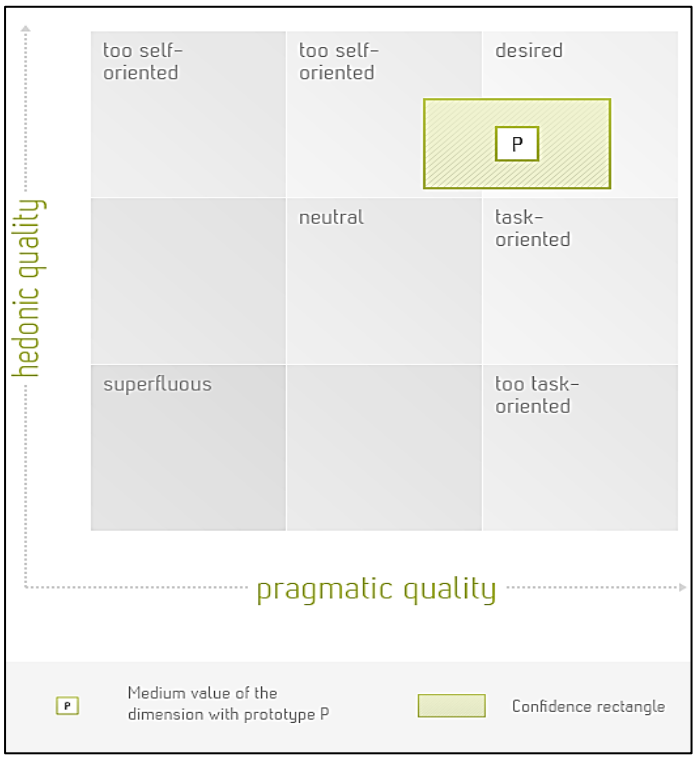



*Figure 5. Attrakdiff dimensions explanation: The prototype P was rated well in both hedonic and pragmatic quality. There was little room for optimization. The confidence rectangle shows that according to user consensus, the hedonic quality is greater than the pragmatic quality. For prototype P the confidence rectangle extends from the desired area and into the self-oriented area. It can therefore not clearly be classified as desirable.*

The process of using AttrakDiff is straightforward: users are asked to rate the product based on pairs of opposite adjectives (Figure 6), such as "simple - complex" or "attractive - unattractive.". These ratings are then analyzed to reveal users' perceptions of the product. A significant advantage of AttrakDiff is its ability to provide detailed information about users' emotional and functional reactions.


*Figure 6. Rating items in Attrakdiff.*

According to research, AttrakDiff has been widely used and proven to be a reliable and valid tool for UX evaluation. Rauschenberger et al. (2013) note that AttrakDiff can provide valuable insights for improving product design, helping designers better understand users' needs and preferences.

Overall, AttrakDiff is a valuable tool for UX evaluation, providing a comprehensive view of the pragmatic and hedonic quality of products. Its use can significantly contribute to the development of products that not only meet functional requirements but also offer a pleasant and attractive user experience.

In evaluation of learning platform Liapis et al. (2023) employed Attrakdiff in order to evaluate and compare the attractiveness dimensions of two open MOOC platforms (Moodle Vs Open edX).

## 4.1. Attrakdiff - Results

In this section we present result from Attrakdiff questionaire responses. Regardning the placement of ADA platform in the pragmatic and hedonic dimensions. Figure 7 represents the appropriate “portfolio-presentation” chart. The latter depicts the mean value and confidence levels for pragmatic and hedonic qualities of the evaluated platform. The location of the points also

shows in which area (i.e. from too self-oriented to too task-oriented) the platform belongs to. The bigger the confidence rectangle the less sure one can be to which region it belongs. For the ADA platform, the value of pragmatic quality is PQ = 0.77 and hedonic quality HQ=. The bigger the confidence rectangle the less sure one can be to which region it belongs. **Our study showed that ADA platform placed in Neutral region in both Pragmatic and Hedonic dimensios.**

A small confidence rectangle is an advantage because it means that the investigation results are more reliable and less coincidental (Figure 7). The confidence rectangle shows, if the users are at one in their evaluation of the product. The bigger the confidence rectangle, the more variable the evaluation ratings.

Figure 8 illustrates the diagram of average values of all answers. The horizontal axis shows the groups (PQ, HQ, ATT). The vertical axis represents the average values of the word pairs inside each group. A bigger number on the vertical axis should be considered as a better UX, while a value that approximates to 0 expresses a neutral experience. **In our study, users found ADA to be better in terms of the pragmatic quality and attractiveness dimensions**.

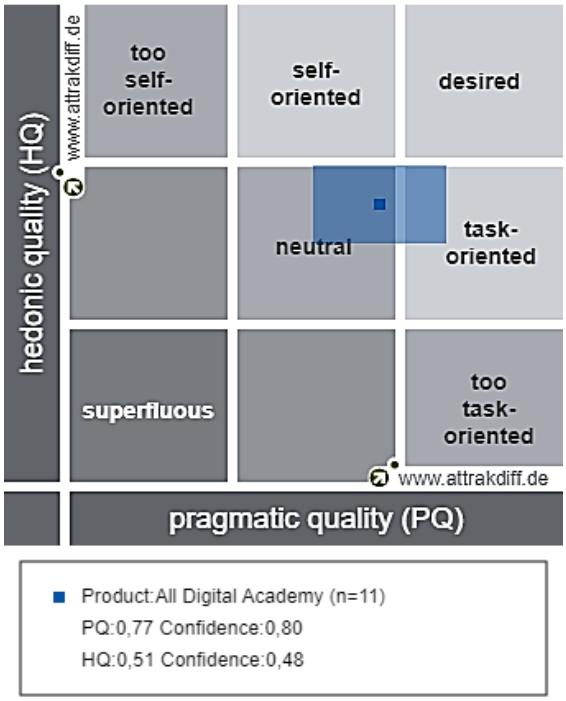



*Figure 7. AttrakDiff portfolio-presentation chart for ADA platform.*

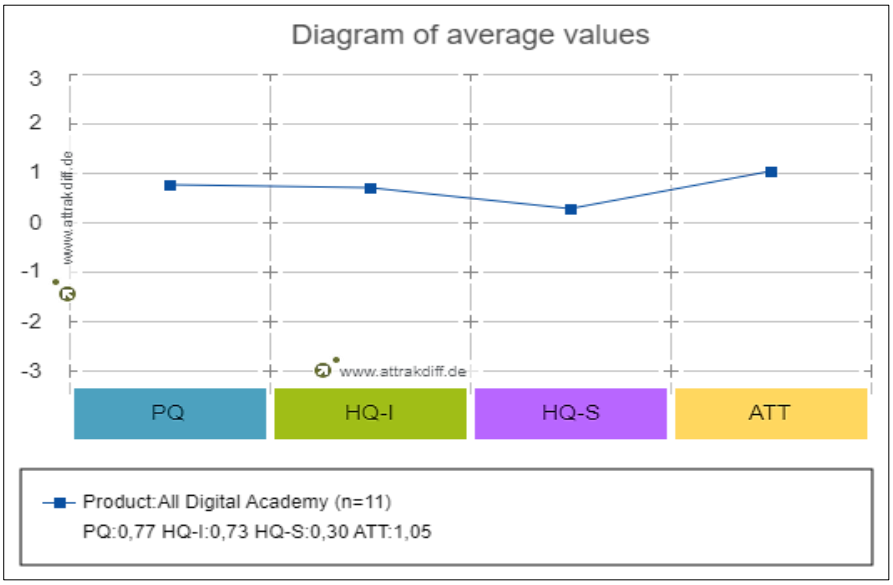



*Figure 8. AttrakDiff diagram of average values (pragmatic quality, hedonic quality, attractiveness) for ADA platform*

The mean values of the word pairs are presented in Figure 9. Of particular interest are the extreme values. These show which characteristics are particularly critical or particularly well-resolved. In our study we did not find any extreme values. One important finding is that platforms attractiveness scores better than other dimensions. Overall, the results show that the product scores

positively across all dimensions, indicating a well-rounded and highly favorable user experience. The product is perceived as practical, stylish, professional, connective, inventive, creative, and attractive, contributing to a high level of overall attractiveness.

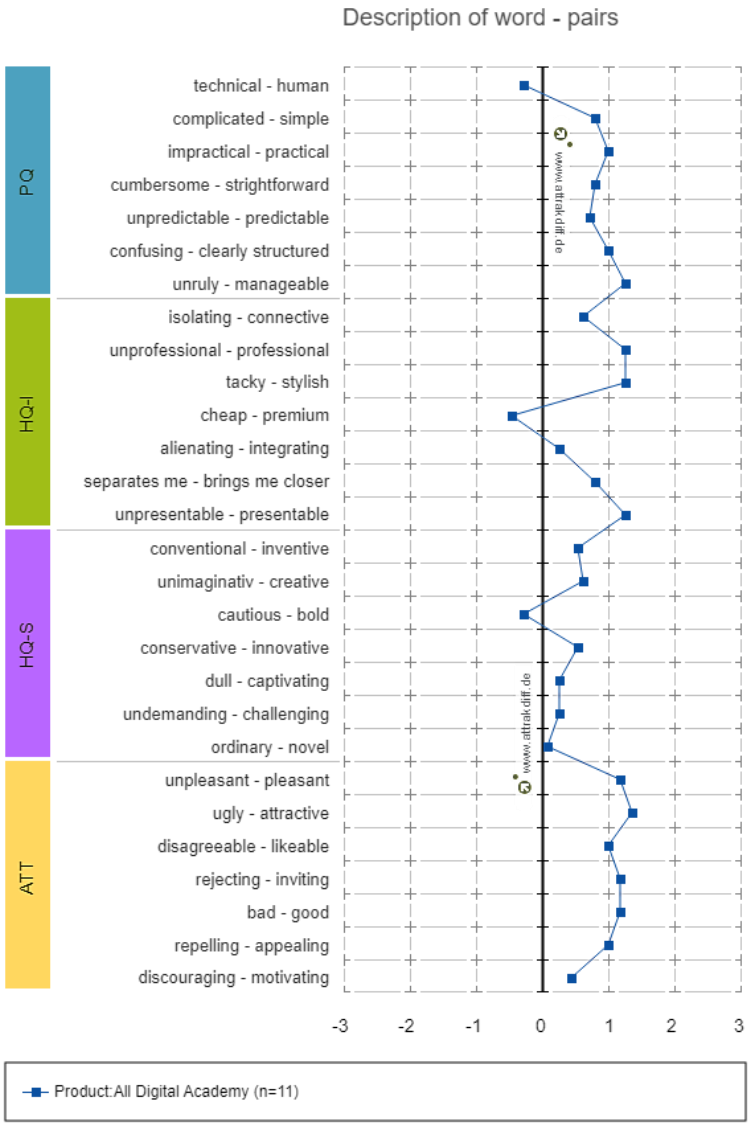



*Figure 9. Description of words – pairs of the AttrakDiff questionnaire for ADA platform.*

Here is a summary of the specific results shown in Figure 9:

**Pragmatic Quality (PQ)**

- **Technical - Human**: The product leans more towards being human than technical.
- **Complicated - Simple**: The product is rated as more simple than complicated.

- **Impractical - Practical**: Users find the product practical.
- **Cumbersome - Straightforward**: The product is considered straightforward rather than cumbersome.
- **Unpredictable - Predictable**: The product is seen as predictable.
- **Confusing - Clearly Structured**: Users perceive the product as clearly structured.
- **Unruly - Manageable**: The product is viewed as manageable.

**Hedonic Quality - Identity (HQ-I)**

- **Isolating - Connective**: The product is seen as more connective than isolating.
- **Unprofessional - Professional**: Users find the product professional.
- **Tacky - Stylish**: The product is rated as stylish.
- **Cheap - Premium**: The product leans towards being premium.
- **Alienating - Integrating**: The product is perceived as integrating.
- **Separates me - Brings me closer**: Users feel the product brings them closer.
- **Unpresentable - Presentable**: The product is considered presentable.

**Hedonic Quality - Stimulation (HQ-S)**

- **Conventional - Inventive**: The product is seen as inventive.
- **Unimaginative - Creative**: Users find the product creative.
- **Cautious - Bold**: The product is rated as bold.
- **Conservative - Innovative**: The product is considered innovative.
- **Dull - Captivating:** Users find the product captivating.
- **Undemanding - Challenging**: The product is perceived as challenging.
- **Ordinary - Novel**: The product is viewed as novel.

**Attractiveness (ATT)**

- **Unpleasant - Pleasant**: The product is rated as pleasant.
- **Ugly - Attractive**: Users find the product attractive.

- **Disagreeable - Likeable**: The product is considered likeable.
- **Rejecting - Inviting**: The product is seen as inviting.
- **Bad - Good**: Users rate the product as good.
- **Repelling - Appealing**: The product is perceived as appealing.
- **Discouraging - Motivating**: The product is seen as motivating.

# 5. Traditional Usability Metrics: time on task and success rate

In the realm of user experience (UX) design, measuring and evaluating usability is paramount for creating effective and user-friendly products. Among the various metrics available, traditional usability metrics such as "time on task" and "success rate" stand out due to their straightforwardness and effectiveness in providing tangible insights into user behavior and system performance. This chapter delves into these fundamental metrics, exploring their significance, application, and impact on usability testing.

Time on task reveals the efficiency of task completion, while success rate demonstrates the effectiveness of the interface. **When used in tandem, they provide a balanced assessment that can drive meaningful enhancements in user experience**. As we explore these metrics in greater detail, we will uncover their practical applications, interpret their results, and discuss how they can be leveraged to create more intuitive and user-friendly designs.

## 5.1. Time on task

Time on task is a critical metric that measures the amount of time a user takes to complete a specific task. This metric is valuable because it provides direct insight into the efficiency of a user interface**. Shorter completion times generally indicate that the system is intuitive and easy to navigate, whereas longer times may suggest complexity or design flaws that hinder user performance**. By analyzing time on task, designers can pinpoint areas where users struggle and make informed decisions to streamline processes, thus enhancing overall user satisfaction.

*Table 3. Descriptive statistics about time on task metric.*

| | **Mean time** | **Std.** | **95% CI[1]** | **95% CI** |
|---|---|---|---|---|

[1] The 95% confidence intervals (CI) provide a range within which we can be 95% confident that the true mean time lies.

| | (in sec.) | Deviation | Lower Bound | Upper Bound |
|---|---|---|---|---|
| ***Task 1*** | 132,61 | 98,42 | 73,14 | 192,09 |
| ***Task 2*** | 166,54 | 108,53 | 100,95 | 232,12 |
| ***Task 3*** | 125,23 | 100,96 | 64,22 | 186,24 |
| ***Task 4*** | 117,46 | 91.47 | 62.18 | 172,74 |
| ***Task 5*** | 96,54 | 102,94 | 34,33 | 158,75 |

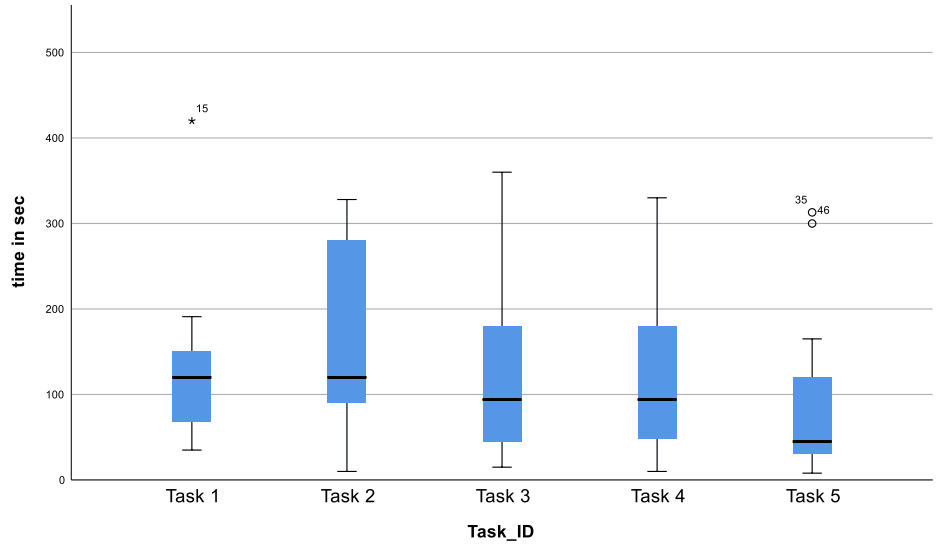



*Figure 10. Time on task box plots. Stars (*) indicates extreme values and (O) indicates outliers.*

The data reveals that Task 2 is the most time-consuming and has the highest variability, suggesting it may be the most complex or least intuitive task. Task 5, while having the shortest mean time, also shows significant variability, indicating differing user experiences. Overall, these metrics highlight areas where usability improvements can be targeted, especially in reducing time on task and standard deviation to achieve more consistent user performance.

### 5.1.1. Insights from the Analysis

**Task 2** has the highest mean time (166,54 sec), indicating it may be the most complex or challenging task. **Task 5** has the shortest mean time (96,54 sec.), suggesting it is the simplest task or the one with the most intuitive design. Regarding SD analysis, **Task 2** also has the highest standard deviation (108,53 sec), indicating a wide variation in user performance. This suggests that some users may find this task significantly more difficult than others. **Task 4** has the

lowest standard deviation (91,47 sec), implying more consistent performance among users. Finally, regarding CI, narrower intervals indicate more precise estimates. **Task 4** has a relatively narrow interval compared to others, suggesting greater precision in the meantime measurement.

## 5.2. Success rate

Success rate measures **the percentage of users who successfully complete a given task without errors**. This metric is a clear indicator of the effectiveness and reliability of a system. A high success rate typically signifies that users can achieve their goals with ease, reflecting well on the usability of the product. Conversely, a low success rate can highlight problematic areas that require attention. Success rate helps in identifying both major and minor usability issues, guiding iterative design improvements to ensure that the system meets user needs effectively.

In general, task success rate was **89,23%**. More specifically, Figure 11 shows the success pes task. A 78% is an acceptable average completion rate (Sauro 2011b).

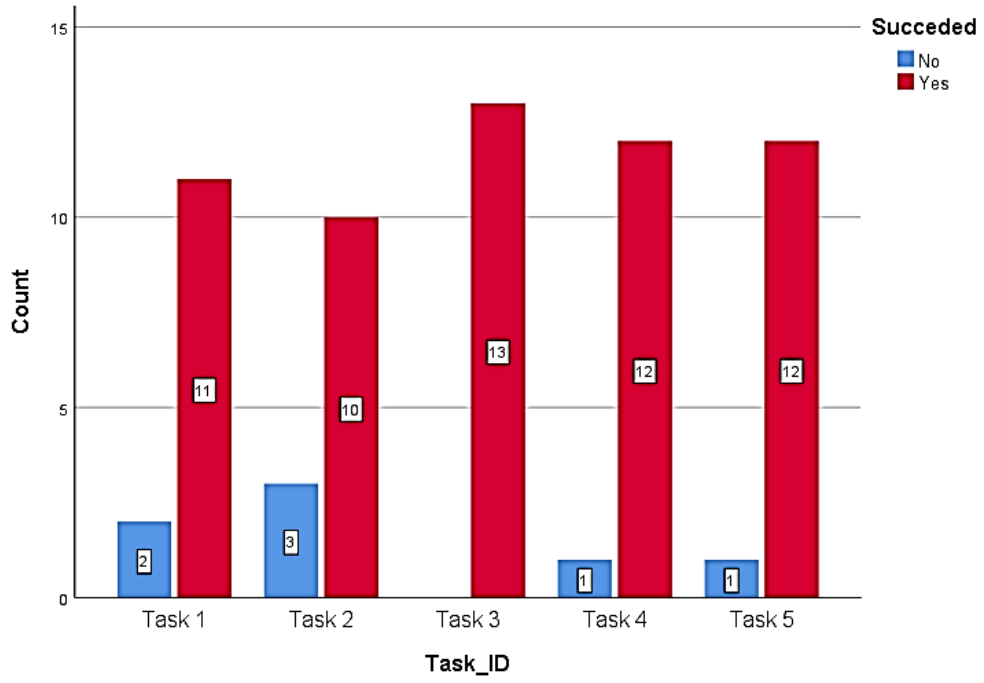



*Figure 11. Success rate per task*

Table 4 present the completion rate per task. Task 2 (Navigate in the platform in order to find and read the research paper entitled "**Artificial Intelligence for HCI: A Modern Approach**") has the lowest completion rate among the five tasks,

suggesting that approximately one in four users struggles to complete it. This indicates a need for usability improvements.

*Table 4. Completion Rate (%) per task.*

| | **Completion Rate (%)** |
|---|---|
| ***Task 1*** | 84,62% |
| ***Task 2*** | 76,92% |
| ***Task 3*** | 100,00% |
| ***Task 4*** | 92,31% |
| ***Task 5*** | 92,31% |

# 6. Beyond User Testing: Usability Evaluation by Experts Using Nielsen's 10 Heuristics

Among the various methods for evaluating usability, expert reviews stand out for their effectiveness and efficiency. One of the most influential frameworks for expert usability evaluation is Jakob Nielsen's 10 usability heuristics. These heuristics are broad principles that provide a structured approach for identifying usability issues in interactive systems.

Nielsen's heuristics, first introduced in 1994, are a set of general rules of thumb that have been widely adopted and validated through extensive use in the field of UX. They serve as a foundational tool for experts to systematically assess the usability of a system, covering essential aspects such as visibility of system status, match between system and the real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimalist design, help users recognise, diagnose, and recover from errors, and help and documentation.

Expert reviews using Nielsen's heuristics are invaluable because they can quickly and efficiently identify usability problems that might not be immediately apparent through user testing alone. Experts applying these heuristics bring their extensive knowledge and experience to the evaluation process, providing insights that can significantly improve the design. This method is particularly beneficial in the early stages of design when iterative testing and refinement can save time and resources by addressing potential issues before they affect a broader user base.

This chapter delves into each of Nielsen's 10 usability heuristics, explaining their significance and application. It also discusses the methodology of conducting

heuristic evaluations, highlighting the role of expert evaluators in identifying and prioritising usability issues. By understanding and applying these heuristics, designers and developers can create more user-friendly interfaces that enhance the overall user experience.

Nielsen's 10 usability heuristics remain a cornerstone of UX design, offering a practical and reliable framework for expert usability evaluations. Their continued relevance and effectiveness underscore their importance in the ever-evolving field of user experience, making them an essential tool for any UX professional aiming to create intuitive and effective digital products.

## 6.1. Conducting a Heuristic Evaluation

In our study, we selected a team of **3 usability experts to perform the evaluation**. The team is experienced with heuristic evaluation and has deep understanding of usability principles. All experts, hold a PhD in the field of Human-Computer Interaction.

Each evaluator independently examined the interface of ADA platfor and evaluateed it against Nielsen's 10 heuristics. They navigate through the platform, performing both predefined and ad hoc tasks in order to note any usability issues they encounter. Evaluators should consider the following rules (Heuristics):

1. **Visibility of system status**: Is the system providing appropriate feedback?
2. **Match between system and the real world**: Does the system speak the users' language?
3. **User control and freedom**: Can users easily undo actions?
4. **Consistency and standards**: Are there any inconsistencies in the interface?
5. **Error prevention**: Does the design prevent potential errors?
6. **Recognition rather than recall**: Are options and information easily visible?
7. **Flexibility and efficiency of use**: Can both novice and experienced users use the system efficiently?
8. **Aesthetic and minimalist design**: Is the design free of unnecessary elements?
9. **Help users recognise, diagnose, and recover from errors**: Are error messages helpful?
10. **Help and documentation**: Is help easily accessible?

## 6.2. Aggregated Results from Expert

Aggregating results from expert evaluations involves compiling and synthesising the findings from each evaluator. This process not only highlights common usability issues identified by multiple experts but also uncovers less obvious problems that may be noted by only one or two evaluators. The aggregation process includes categorising issues, assigning severity ratings, and prioritising them based on their impact on the user experience.

The benefits of this aggregated approach are manifold. Firstly, it minimises the bias that can occur when relying on a single evaluator's perspective, ensuring a more balanced and objective assessment. Secondly, it provides a comprehensive overview of the usability landscape, capturing a wider range of issues that may affect different user groups. Lastly, it helps in prioritising usability improvements, guiding the development team to focus on the most critical issues that, once resolved, will significantly enhance the overall user experience.

In this chapter, we will explore the methodologies for aggregating expert evaluation results, discuss best practices for synthesising and analysing these findings, and provide practical examples of how aggregated data can inform and drive usability improvements. By understanding and implementing effective aggregation techniques, usability professionals can ensure their evaluations lead to meaningful and impactful enhancements in user interface design.

### 6.2.1. Main Findings

**Issue-1:** After logging in, users are directed to the Activity Dashboard. However, if users select "Home" from the navigation, they are redirected to the initial landing page seen before logging in. **This action is confusing because the main navigation menu disappears**, and users must select "Login," "Register," or use the browser's "Back" button to return to the Activity Dashboard or other parts of the site. **This disrupts the user experience and can cause frustration** as users struggle to navigate back to their desired location.

**Recommendations:** To enhance usability and reduce confusion**, the "Home" option should be disabled or hidden while users are logged in**. Instead, ensure that logged-in users remain within the authenticated area of the site, such as the Activity Dashboard or other relevant sections. This approach maintains a consistent navigation experience, helping users to easily find their way and understand where they are within the application**.**

**Violation**: The primary heuristic violated by this issue is **User Control and Freedom**, as it restricts users' ability to navigate the site efficiently and recover from their actions. Additionally, it also touches on **Match Between System and the Real World** and **Visibility of System Status**, which contribute to the overall confusion and disorientation experienced by users.

Rating: **3 (Major usability problem)**. **It is important to fix this issue to ensure a smooth and intuitive user experience, preventing unnecessary frustration and confusion for users**. This should be given high priority in the usability improvement process.

---

**Issue-2:** Clicking the FAQ menu redirects users to a **'coming soon**' page instead of the expected FAQ content. This creates confusion and hinders users from finding answers to their questions.

**Recommendations: It is recommended to address this issue promptly**. Ensure that the FAQ menu accurately reflects the current status of the content. If the FAQ section is not yet available, consider providing a temporary solution such as a message indicating when the content will be accessible. Alternatively, redirect users to a page that provides related information or directs them to contact support for assistance.

**Violation**: **3 (Major usability problem)**. Important to fix and should be given high priority.

---

**Issue-3**: In contact form the link “Terms of use & the Privacy policy.” Is not working. More specifically, it returns 404 error (page not found).

**Recommendations**: Double-check the URL configured for the "Terms of Use & Privacy Policy" link in the contact form. Ensure that it is correctly pointing to the actual location of these documents on your website or server.

**Violation**: The issue where the "Terms of Use & Privacy Policy" link in the contact form returns a 404 error typically violates several usability heuristics, primarily related to **visibility of system status** and **error prevention.**

**Rating**: **2 (Minor usability problem)**. Fixing this should be given low priority.

---

**Issue-4**: NEWS and EVENTS sections on the ALL DIGITAL Academy platform require a lot of scrolling to view items.

**Recommendations**: **Layout might benefit from pagination or a condensed view.** This would make it easier for users to access and navigate through the information without excessive scrolling. Implementing such features can improve user experience by allowing quick access to specific articles or events. Provide users with options to filter news and events based on categories, dates, or keywords. This allows users to narrow down the list to find items of interest more quickly, reducing the need for extensive scrolling.

**Violation**: **Efficiency of use**: Users should be able to accomplish tasks quickly and with minimal effort. Excessive scrolling to access news and events can frustrate users and make it difficult for them to find relevant information efficiently.

**Rating**: **2 (Minor usability problem)**. Fixing this should be given low priority.

---

**Issue-5**: Accesibility functionalities are not available by default. Users must activate them after an extensive navigation in platform. More specifically, accesibility tools are available in users' profile - > My MOOCs. Next, users have to select the new profile menu (Figure 12). A new menu appears and users can select accesibility.

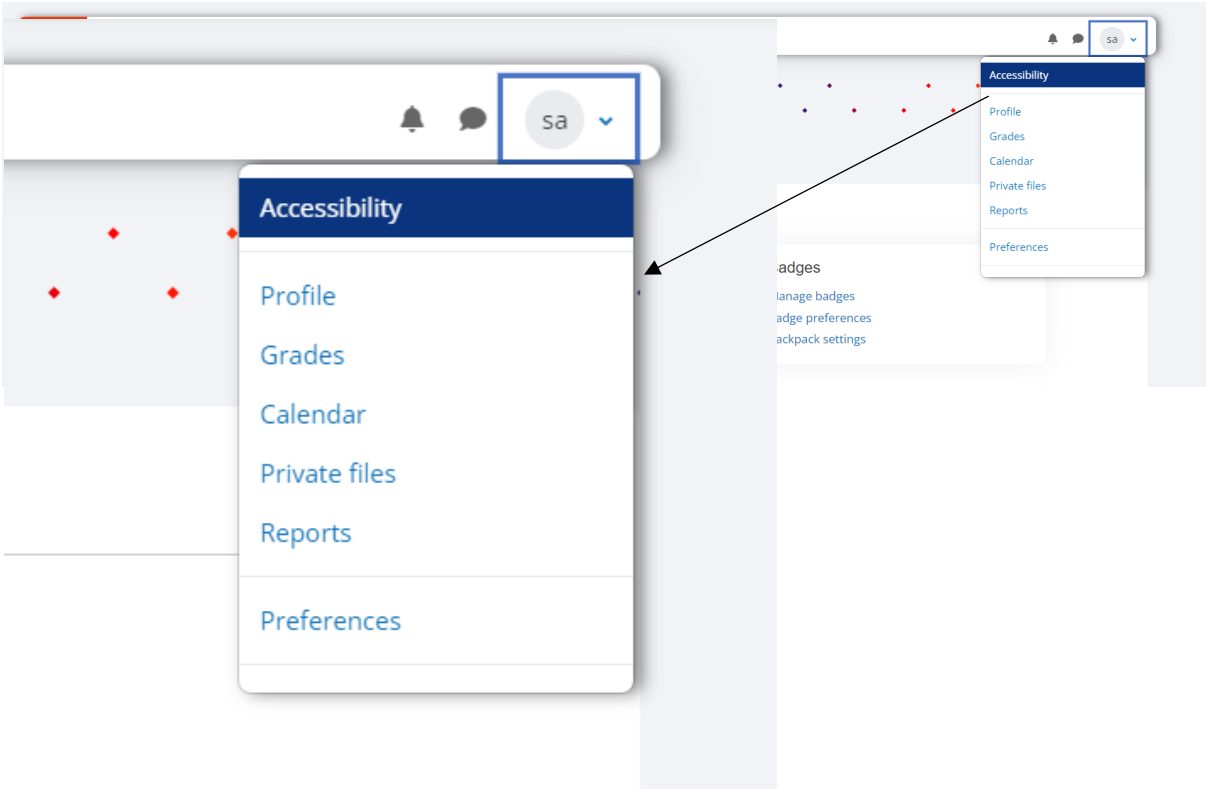



*Figure 12. Activation of Accessibility functionality.*

**Recommendations:** Consider **placing accessibility options in a more visible and easily accessible location**, such as within the main navigation menu or in a dedicated section in the footer of the website. This ensures users can find and activate them without extensive searching.

**Violations**: This issue violates several usability principles, particularly those related to accessibility and user control. More specifically, the issue violates:

a. **Visibility of system status**: Users should be aware of the system's current state and available functionalities without having to search extensively. If accessibility tools are hidden deep within the user profile settings, it lacks clear visibility of their availability
b. **User control and freedom**: Users should have the freedom to access important functionalities, such as accessibility tools, easily and intuitively. Requiring users to navigate through multiple menus and options to find accessibility settings restricts their control over how they interact with the platform
c. **Consistency and standards**: Accessibility features should ideally be easily accessible and follow standard conventions. Placing them deep within profile settings rather than in a more prominent or universally accessible location goes against standard practices for making platforms inclusive.

**Rating**: **3 (Major usability problem)**. Significantly affects usability and accessibility, particularly for users who depend on these features.

**Issue-6**: In the message sending form the reset option clears subject and body, and the username of the recipient. Regarding the latter, this is something that should not be happening.

**Recommendations:** The specific action should be redesigned. More specifically, if a user wants to delete a recipient an appropriate symbol option (e.g. recycle bin, minus symbol) next to recipient must be employed. **Furthermore, implement a confirmation prompt or dialogue box when the reset button is clicked, especially if it's intended to clear multiple fields**. This allows users to confirm their intention and prevents accidental data loss.

**Violations: Error prevention**: Users should be able to recover from errors easily and should not be penalised for using system functionalities such as a reset button. Clearing the recipient's username without warning or confirmation can lead to user frustration and potential data loss, especially if the username was entered correctly and accidentally reset.

**Rating**: **3 (Major usability problem)**. Significantly affects usability and accessibility, Important to fix, and should be given high priority.

---

**Issue-7:** If a user tries to send a Private Message without putting subject and/or body, the platform correctly displays an appropriate error message. If the user fills in what was missing, the message is not sent, and user should reload the page and do the steps from the beginning.

**Recommendations:** Implement client-side form validation that checks for missing subject and/or body fields before the form is submitted. When an error occurs (e.g., missing fields), display an inline error message next to the relevant fields and allow users to correct them without clearing other inputs.

**Violations**: **Error recovery**: Users should be able to recover from errors easily and continue their tasks without unnecessary friction. In this case, after receiving an error message about missing subject and/or body fields, users

should be able to correct these fields and proceed with sending the message without having to reload the page or start over from the beginning.

**Rating**: **3 (Major usability problem)**. Significantly affects usability and accessibility, Important to fix, and should be given high priority.

## References

Albert, William, and Thomas Tullis. 2013. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Newnes.

Anyatasia, F. N., H. B. Santoso, and K. Junus. 2020. "An Evaluation of the Udacity MOOC Based on Instructional and Interface Design Principles." *Journal of Physics: Conference Series* 1566(1):012053. doi: 10.1088/1742-6596/1566/1/012053.

Bangor, Aaron, Philip T. Kortum, and James T. Miller. 2008. "An Empirical Evaluation of the System Usability Scale." *International Journal of Human–Computer Interaction* 24(6):574–94. doi: 10.1080/10447310802205776.

Bevan, Nigel. 2009. "What Is the Difference between the Purpose of Usability and User Experience Evaluation Methods?"

Brooke, John. 1996. "SUS-A Quick and Dirty Usability Scale." *Usability Evaluation in Industry* 189(194):4–7.

Chin, John P., Virginia A. Diehl, and Kent L. Norman. 1988. "Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface." Pp. 213–18 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *CHI '88*. New York, NY, USA: ACM.

Hassenzahl, Marc, Michael Burmester, and Franz Koller. 2003. "AttrakDiff: A Questionnaire to Measure Perceived Hedonic and Pragmatic Quality." Pp. 187–96 in Vol. 57. Springer Heifelberg.

Hassenzahl, Marc, and Noam Tractinsky. 2006. "User Experience - a Research Agenda." *Behaviour & Information Technology* 25(2):91–97. doi: 10.1080/01449290500330331.

Hornbæk, Kasper, and Effie Lai-Chong Law. 2007. "Meta-Analysis of Correlations among Usability Measures." Pp. 617–26 in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM.

Korableva, Olga, Thomas Durand, Olga Kalimullina, and Irina Stepanova. 2019. "Studying User Satisfaction with the MOOC Platform Interfaces Using the Example of Coursera and Open Education Platforms." Pp. 26–30 in *Proceedings of the 2019 International Conference on Big Data and Education*, *ICBDE '19*. New York, NY, USA: Association for Computing Machinery.

Lewis, James R., and Jeff Sauro. 2009. "The Factor Structure of the System Usability Scale." Pp. 94–103 in. Springer.

Liapis, Alexandros, Vicky Maratou, Theodor Panagiotakopoulos, Christos Katsanos, and Achilles Kameas. 2023. "UX Evaluation of Open MOOC Platforms: A Comparative Study between Moodle and Open edX Combining User Interaction Metrics and Wearable Biosensors." *Interactive Learning Environments* 31(10):6841–55.

Norman, Don, and Jakob Nielsen. 2014. "The Definition of User Experience (UX)." *Nielsen Norman Group*. Retrieved (https://www.nngroup.com/articles/definition-user-experience/).

O'Brien, Heather L., and Elaine G. Toms. 2010. "Is There a Universal Instrument for Measuring Interactive Information Retrieval?: The Case of the User Engagement Scale." Pp. 335–40 in *Proceedings of the Third Symposium on Information Interaction in Context*, *IIiX '10*. New York, NY, USA: ACM.

Rauschenberger, Maria, Martin Schrepp, Manuel Pérez Cota, Siegfried Olschner, and Jörg Thomaschewski. 2013. "Efficient Measurement of the User Experience of Interactive Products. How to Use the User Experience Questionnaire (UEQ). Example: Spanish Language Version."

Sauro, Jeff. 2011a. *A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices*. Measuring Usability LLC.

Sauro, Jeff. 2011b. "What Is a Good Task-Completion Rate." *Retrieved June* 1:2022.

Stetson, Jacqueline N., and Thomas S. Tullis. 2004. “A Comparison of Questionnaires for Assessing Website Usability.” *UPA Presentation*.

Tsironis, A., C. Katsanos, and M. Xenos. 2016. “Comparative Usability Evaluation of Three Popular MOOC Platforms.” Pp. 608–12 in *2016 IEEE Global Engineering Education Conference (EDUCON)*.

Vercellotti, Mary Lou. 2018. “Do Interactive Learning Spaces Increase Student Achievement? A Comparison of Classroom Context.” *Active Learning in Higher Education* 19(3):197–210.

Vlachogianni, Prokopia, and Nikolaos Tselios. 2021. “Investigating the Impact of Personality Traits on Perceived Usability Evaluation of E-Learning Platforms.” *Interactive Technology and Smart Education* 19(2):202–21. doi: 10.1108/ITSE-02-2021-0024.

Wechsung, Ina, Klaus-Peter Engelbrecht, Christine Kühnel, Sebastian Möller, and Benjamin Weiss. 2011. “Measuring the Quality of Service and Quality of Experience of Multimodal Human–Machine Interaction.” *Journal on Multimodal User Interfaces* 6:73–85. doi: 10.1007/s12193-011-0088-y.

Zaharias, Panagiotis, and Angeliki Poylymenakou. 2009. “Developing a Usability Evaluation Method for E-Learning Applications: Beyond Functional Usability.” *International Journal of Human–Computer Interaction* 25(1):75–98. doi: 10.1080/10447310802546716.