Assessing the Usability of Federated Access to T4EU Online Courses in Higher Education Mobility Programs

Federica Mancini[1], Riccardo Fattorini, Michele Bava

University of Trieste – Trieste (Italy)

Abstract

Facilitating access to online courses in higher education mobility programs is essential for creating a more interconnected educational ecosystem within the European Education Area. Federated e-infrastructures have emerged as effective solutions to enhance the interoperability, accessibility, and scalability of academic services under a standardized trust model. However, assessing their usability for end-users is critical. This study aims to identify and adapt an instrument for measuring the usability of federated access to a Moodle ecosystem implemented by the Transform4Europe alliance for students participating in mobility programs. The paper outlines the process of adapting and validating a questionnaire based on Nielsen's Usability Attributes model to meet the unique characteristics of this context. An iterative, multi-method approach was employed, incorporating feedback from students and usability experts for content validation. The resulting instrument was administered to 145 students at the University of Trieste during lectures. Exploratory factor analysis confirmed the tool's reliability and validity while highlighting the need for refinements, including revising two items with low factor loadings, methodological adjustments in questionnaire administration, and increased sample size for more robust results. Although further validation of the final instrument is recommended, the results obtained in this study provide a significant starting point for advancing usability assessment practices in federated learning environments aimed at enhancing the student mobility experience.

KEYWORDS: European Universities, e-Learning, Federated Access, Moodle, Online Courses, Usability.

1. Introduction

The digital transformation of university campuses and the increasing adoption of distance learning, supported by European Universities Alliances (EUAs), are essential for creating a more interconnected and student-centered educational ecosystem in the European Higher Education Area (EHEA) (Gaebel et al., 2021). However, developing joint campuses and exchanging electronic data across systems remains challenging for Higher Education Institutions (Berger et al., 2023).

To address the need for a global learning environment, the Transform4Europe (T4EU) alliance implemented a Moodle ecosystem accessible exclusively to alliance members through federated authentication, leveraging the eduGAIN inter-federation service. After the testing phase, we evaluated the usability of this federated ecosystem based on students' experiences at the University of Trieste (UniTS), before extending the study to other institutions. This evaluation involved identifying a usability instrument tailored to the specific context.

Usability, as defined by ISO 9241-11, refers to the degree to which a user can utilize a product to achieve specific goals with effectiveness, efficiency, and satisfaction within a defined context of use (ISO, 2018). Usability evaluation focuses on users' ability to learn and use a product to accomplish their objectives and the satisfaction they experience during use. Several methods can be used for such evaluations, with usability questionnaires being a widely adopted and cost-effective option that provides valuable insights

---

[1] corresponding author - email: federica.mancini@units.it

into user perceptions (Aziz et al., 2021). However, selecting the most appropriate usability questionnaire can be challenging.

The literature offers limited guidance on the best questionnaire for evaluating usability in federated Moodle ecosystems for mobility students. For example, Ruoti et al. (2015) use the Systems Usability Scale (SUS) to assess the usability of web authentication systems, confirming its reliability. Galende et al. (2023) and Vlachogianni and Tselios (2023) also use SUS, highlighting its widespread adoption for perceived usability evaluation in educational platforms.

Despite the popularity of questionnaires, Sagar and Saha (2017) found no consensus on the models used for usability analysis. Hodrien and Fernando (2021) suggest selecting the right instrument involves analysing the study context and systems and evaluating the questionnaire's content, advantages, disadvantages, and psychometric properties. The questionnaire should also be easy to administer and adaptable to the context.

Considering this scenario, the main research questions guiding this research are:

R1. How can the usability of federated access for online courses in higher education mobility programs be accurately evaluated using a questionnaire?

R2. How does the re-adapted Nielsen Attributes of Usability (NAU)-based questionnaire measure the usability of federated access to online courses in higher education mobility programs?

2. Materials and methods

2.1 Selection and Adaptation of the Instrument

Inquiry methods for collecting quantitative data from both students and experts were crucial for our study. Selecting the most appropriate questionnaire required a thorough review of widely used usability instruments (Brooke, 1996; Chin et al., 1988; Kirakowski & Cierlik, 1998; Kirakowski, 1995; Laugwitz et al., 2008; Lewis, 1992; Lund, 2001; Nielsen & Kaufmann, 1993). The research team assessed the items based on their relevance to our context, where navigation spans multiple screens and services rather than a single system. Furthermore, attention was given to the number of items, as the instrument needed to be completed quickly during lectures, while also capturing both utilitarian (performance-focused) and experiential (satisfaction-focused) aspects (Chung & Sahari, 2015).

Each researcher independently evaluated the questionnaires, after which the team discussed the results to reach a consensus. This process led to the selection of Nielsen's Usability Attributes (NAU) (Nielsen & Kaufmann, 1993), a flexible framework featuring five customizable attributes in a concise format. Additionally, NAU assesses both functional and experiential aspects, making it well-suited for

evaluating federated access systems where functionality, security, and privacy are critical.

Although various NAU-based questionnaires exist in the literature (Benmoussa et al., 2019; Gonzalez-Holland et al., 2017; Halim et al., 2021; Latiar et al., 2024; Munir et al., 2019), none have been formally validated. Furthermore, while NAU attributes have been translated into multiple languages, no validated Italian version was available. For this study, we adopted Benmoussa's version and translated it into Italian. Two independent translators worked on the translation, and the drafts were merged by the research team to ensure consistency and accuracy. The term "system" was replaced with "procedure" to better suit the evaluation context.

A multi-method approach (Palmieri et al., 2020) was adopted for iterative data collection and adjustments over three phases, refining the instrument for optimal use.

2.2 Phase 1: Content Validation with Students

After completing the basic adaptation and translation, a content validity analysis was conducted to evaluate the instrument's coverage of usability domains (Bandalos, 2018) and to eliminate irrelevant items (Boudreau et al., 2001; Lewis et al., 1995). Additionally, the language was reviewed for clarity to ensure readability and comprehension.

Between September and December 2023, ten UniTS students participating in the T4EU mobility programs tested the federated access procedures for Moodle courses at four alliance institutions: the University of Alicante (UA), the University of Silesia (USil), Sofia University 'St. Kliment Ohridski' (SU), and Vytautas Magnus University (VMU). These universities were chosen for the pilot due to their successful implementation of federated authentication through EduGain. As such, five students tested each procedure, a sample size deemed sufficient to uncover approximately 85% of usability issues (Nielsen, 2012). The testing was conducted in moderated, face-to-face sessions, which were recorded for subsequent analysis. Usability feedback was collected in three stages:

- First Procedure Test and Overall Clarity Feedback: After testing the first federated access procedure, students completed the usability questionnaire and provided feedback on its overall clarity.
- Second Procedure Test and Item Clarity Feedback: In the second stage, students tested the procedure at a second institution, re-completed the questionnaire, and rated the clarity of each item on a five-point Likert scale. The moderator conducted cognitive interviews for items that received a score below 3.
- Coverage and Relevance Assessment: Students answered targeted questions to evaluate the relevance and coverage of the questionnaire items.

Recordings were qualitatively analysed using Atlas.ti. Two researchers independently coded the data, generating keywords a posteriori (Creswell, 2013). Inter-rater reliability reached 80%. Insights from this process informed a refined second version of the questionnaire.

2.3 Phase 2: Content Validation with Experts

The revised questionnaire was evaluated by ten UniTS usability experts, all with a minimum B2 level of English. Each expert tested one federated access procedure and assessed the questionnaire based on two criteria:

Language Clarity: Experts used a dichotomous scale ("Yes" for clear, "No" for unclear), providing justifications and reformulations for unclear items (Taherdoost, 2016). Clarity scores were averaged across all items to provide an overall assessment.

Content Validity: Experts rated item relevance using a 4-point Likert scale (Lawshe, 1975). Scores of 1-2 were considered irrelevant, while 3-4 were deemed relevant. The Content Validity Ratio (CVR) was calculated using Lawshe's formula, with items scoring ≥ 0.78 regarded as valid (Wilson et al., 2012). The Content Validity Index (CVI) was also calculated to assess the validity of the items.

Experts also recommended additional items to more effectively capture the user experience. Items with high CVR and CVI scores were either retained or refined for the development of the third version of the instrument.

2.4 Phase 3: Construct Validation with Students

Construct validity was evaluated through factor analysis (Strauss & Smith, 2009).

From March to April 2024, the revised questionnaire was administered to 145 UniTS students enrolled in *Computer Literacy*, *Introduction to Sports Psychology*, and *Perception*. Students tested one of four federated access procedures during face-to-face lectures, with absent participants completing the questionnaire via email.

Data analysis was performed using Jamovi 2.3.28. Confirmatory Factor Analysis (CFA) assessed the model fit of the 10-item instrument, applying absolute, incremental, and parsimonious fit indices. Exploratory Factor Analysis (EFA) was then conducted on a 12-item version to examine its data structure (Harman, 1976; Polit & Beck, 2006). EFA suitability was verified using the Kaiser-Meyer-Olkin (KMO) and Bartlett's test.

3. Results

3.1 Phase 1

The translation and adaptation of Benmoussa's (2019) NAU-based questionnaire resulted in the first version of the quesionnaire (VERSION 1), as detailed in Table 1A of the Appendix. The test results show that students consistently rated the language clarity highly, indicating effective comprehension of the items (Fig. 1 of the Appendix).

Despite the strong endorsement of language clarity, this finding is somewhat undermined by the user experience analysis and the overall feedback from participants during the test sessions. While only four items (2, 4, 6, and 10) received low clarity ratings from a single student, nearly all participants raised concerns. Specifically, the terms "procedure" and "screen" were frequently questioned. As students navigated multiple interfaces – moving from the partner university's Moodle login page to the home university for attribute authorization and back to Moodle to access the course – the term "procedure" failed to capture the full navigation path. This led to some confusion, with students uncertain whether they were evaluating the login process or the broader navigation within the LMS to reach the course. Similarly, "screen" was unclear, with participants unsure whether it referred to the login or course display.

The content validity analysis revealed a need for additional context-specific items. Some students suggested including items to assess navigation tools, layout clarity, translations, the logout confirmation message, and course registration details. For instance, one student recommended evaluating the interface's ability to inform users about their navigation context, while another noted the absence of an item clarifying the enrolment process (e.g., whether students are automatically enrolled or need to take action). Additionally, feedback from the Likert scale revealed the absence of a "not applicable" option for items related to errors (items 7 and 8), as some students were unable to provide a rating when no errors occurred.

Based on these findings, the research team revised the questionnaire, resulting in a second version (VERSION 2). To improve clarity, terms like "procedure" and "screen" were replaced with "navigation". It was clarified that "navigation" and "procedure", when retained, referred to the entire user journey, from federated login to course access. The questionnaire was also updated to cover all relevant domains, introducing the attribute 'Effectiveness' with two new items focused on evaluating information quality, including service descriptions and data privacy management during login (see Table 1A of the Appendix). Additionally, the Likert scale was modified to include a "not applicable" option, enhancing the tool's ability to capture the full range of user experiences.

3.2. Phase 2

The results demonstrate strong consensus among usability experts regarding the clarity of the items. Each item was evaluated on a scale where 1 indicates "clear" and 0 represents "unclear." The average score for each item was calculated from these ratings, offering a comprehensive assessment of clarity based on expert feedback (see Table 1).

Table 1 - Language Clarity.

	ITEMS	*Proportion of experts rating as clear*
Learnability	ITEM 1	1
	ITEM 2	1
Efficiency	ITEM 3	1
	ITEM 4	1
Effectiveness	ITEM 5	1
	ITEM 6	1
Memorability	ITEM 7	1
	ITEM 8	1
Errors	ITEM 9	1
	ITEM 10	1
Satisfaction	ITEM 11	1
	ITEM 12	0.9

The findings regarding item relevance demonstrate that each item meets the established threshold, validating their inclusion in the final instrument (see Table 2A of the Appendix).

The content validity index (CVI) for the entire instrument was accurately calculated as the average of the CVR values for all items that reached the threshold of 0.78 and were retained in the final version (Kaiser, 1970). With a CVI score of 0.91, the instrument exhibits outstanding content validity, confirming its efficacy in assessing usability.

3.3. Phase 3

A total of 145 students completed the second version of the questionnaire after testing the access procedures for their assigned institutions, with responses detailed in Table 2. Five outliers were removed due to mean scores below 1 or above 3.90 (specifically 0.00, 0.75, 0.75, 0.92, and 4.00).

An ANOVA was conducted to assess potential differences in mean scores across institutions, which could indicate environmental impacts on the user experience. However, results showed no significant differences, with mean scores ranging from 2.44 to 2.66 (see Table 3A of the Appendix; see Figure 1),

suggesting a consistent user experience across the institutions' access procedures.

Table 2 - Number of tests of each access procedure.

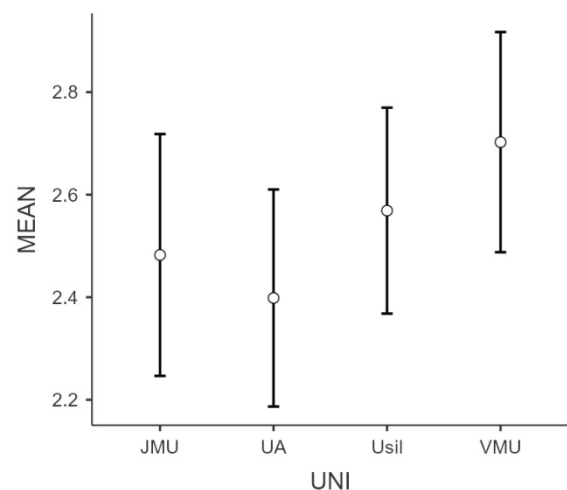| INSTITUTIONS	TESTS
UA	38
USil	42
VMU	36
JMU	29
TOTAL	145

Figure 1 - Graphic representation of the Anova.

A Confirmatory Factor Analysis (CFA) was performed on the data, excluding the two items related to the 'effectiveness' attribute. The objective was to evaluate how well the original five-dimensional NAU model aligned with the data and to determine whether the inclusion of an additional dimension was truly necessary. Absolute fit indices, including RMSEA (0.129) and SRMR (0.0729), indicated poor model fit, as values typically should be below 0.10 and 0.05, respectively (Browne & Cudeck, 1989; Creswell, 2013; Hu & Bentler, 1995; Miles & Huberman, 1994). Incremental fit indices, such as the CFI (0.861) and TLI (0.821), were below the acceptable threshold of 0.90 (Bentler, 1990; Browne & Cudeck, 1989). The parsimonious fit index, derived from the normalized chi-square (χ^2 = 116) and Degrees of Freedom (DF = 35), gave a value of 3.31, indicating a slightly inadequate fit (Cole, 1987; Schermelleh-Engel et al., 2003).

These results showed that the five-dimensional NAU model did not fit well, prompting an Exploratory Factor

Analysis (EFA) that included items 5 and 6 from the new 'Effectiveness' dimension. Factor analysis suitability was confirmed with a KMO value above 0.80, indicating sufficient sample size for the EFA (Polit & Beck, 2006; Harman, 1976) (see Table 3).

Bartlett's sphericity test produced a p-value below 0.05, confirming significant relationships among the items and supporting the factorability of the variables (see Table 4).

Table 3 - KMO Sampling Adequacy Measure.

	MSA
Global	0.868
ITEM 1	0.859
ITEM 2	0.880
ITEM 3	0.925
ITEM 4	0.784
ITEM 5	0.871
ITEM 6	0.898
ITEM 7	0.892
ITEM 8	0.864
ITEM 9	0.773
ITEM 10	0.844
ITEM 11	0.868
ITEM 12	0.904

Table 4 - Bartlett's Test of Sphericity.

χ^2	df	p
695	66	$< .001$

EFA was performed using the 'maximum likelihood' extraction method with 'oblimin' rotation. Before this, skewness and kurtosis analysis confirmed the normal distribution of the data (see Table 4A of the Appendix). Oblique rotation was chosen to account for inter-factor relationships, providing a more accurate representation of the theoretical constructs. Factor loadings of 0.4 or higher were considered significant.

The responses to items 9 and 10, where no error occurred, were treated as missing data in the EFA to maintain the integrity of the dataset and avoid the exclusion of incomplete cases (Bentler & Mooijaart, 1989). Removing these responses could introduce biases and result in the loss of valuable information. To mitigate this risk and in accordance with 'Guideline 1: Use All Available Data' (Newman, 2009, 2014), all 140 cases were included in the analysis. The results confirmed that six factors should be retained (see Table 5A in the Appendix).

The scree plot (Figure 2) shows the relationship between the number of factors and their eigenvalues.
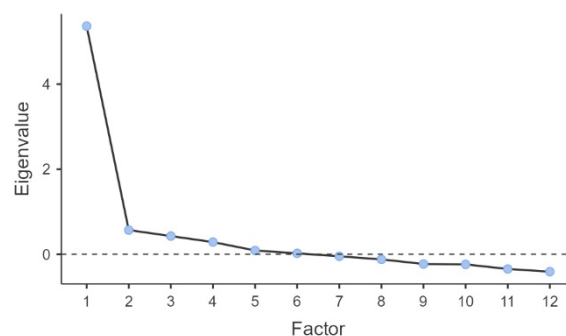

Figure 2 - Scree plot.

The first six factors collectively explain 72.8% of the total variance (see Table 6A in the Appendix), which falls within the commonly accepted range of 70% to 80% for factor extraction (Geisen & Bergstrom, 2017), although according to the Kaiser–Guttman rule (Kaiser, 1960), only the first factor shows an eigenvalue greater than 1. The results also indicate that Items 3 and 12 should be removed, as their factor loadings fall below the acceptable threshold of 0.4.

4. Discussion

The process of selecting the most appropriate instrument for assessing the usability of federated access in online courses within higher education mobility programs has proven highly effective for our purposes, given the specific characteristics of the scenario. A key takeaway is the critical importance of adopting a rigorous, context-driven approach when choosing a usability instrument.

Indeed, a thorough analysis of the Moodle ecosystem, revealed several crucial considerations. For instance, federated access to online courses involves a relatively straightforward procedure consisting of sequential actions, which users are likely to repeat across various local Moodle environments. Navigation spans multiple screens and services rather than being confined to a single system. Furthermore, the authentication and access processes are governed by the eduGain Policies Framework, which necessitates addressing critical issues such as informed consent, cookie usage, and compliance with privacy regulations—factors that significantly influence the navigation and the overall user experience.

As such, the usability evaluation of utilitarian aspects that are linked to the functionality, security, and privacy compliance plays a fundamental role in federated access systems. Given the routine nature of the tasks and their sequential nature, efficiency also emerged as a pivotal usability criterion. Measuring these aspects is therefore a priority over more "hedonic" aspects, which instead concern the pleasure, engagement, and emotional satisfaction derived from interacting with the

system. The analysis also underscored the need for a concise usability instrument with a limited number of items to ensure the evaluation remains both comprehensive and manageable. These insights helped address the first research question, ultimately leading to the selection of the NAU questionnaire as the most suitable instrument for measuring the usability of federated access in online courses after a deliberative process involving consensus among the researchers.

The content validation results also highlighted the need for a change in the test administration method. Given that students perceived authentication and course access as distinct phases, future studies could improve usability assessments by gathering feedback at two separate points in the process: once immediately after federated access and again after course access (see Table 7A in the Appendix). Adopting a 'usability testing with prompt' approach (Hair et al., 2010; Lazar et al., 2017; Shneiderman et al., 2017), supported by specialized software (e.g., UserTesting, Maze, Lookback, Hotjar, UXTweak), would allow questions to be asked at critical navigation points, providing real-time feedback and clearer insights into user interactions and decision-making.

The assessment of the instrument's effectiveness in measuring the usability of federated access to online courses in higher education mobility programs was crucial for addressing the second research question. Furthermore, regarding the validation of the six-factor solution — Learning, Efficiency, Memorability, Error, Satisfaction, and Effectiveness (see Tables 1 and 9 in the Appendix) although the Kaiser criterion was not met for five of the factors, retaining all six was considered appropriate. This decision was supported by the scree plot analysis, which revealed an elbow indicative of a multifactorial solution, and by the cumulative variance explained which meets the commonly accepted threshold for adequately representing the latent structure.

In line with the ISO 9241-11 standard (ISO, 1998), 'effectiveness' emerged as a crucial attribute in measuring usability, particularly in online or hybrid mobility contexts. The cognitive interviews also underscored the importance of clear instructions and timely responses in effectively guiding users through the federated access. Given the complexity and unfamiliarity of the procedure for students, the information provided should not only address basic navigation but also offer comprehensive support that spans the entire virtual mobility experience, ensuring students feel confident and informed at every stage.

The findings also indicated that to improve model accuracy, only 10 of the original 12 items should be retained. Items 3 and 12 were found to inadequately capture key aspects, due to low factor loadings, and require rewording to reduce misinterpretation.

Item 3, "I easily found the information I was looking for," under 'Efficiency,' originally from the NAU questionnaire (Benmoussa et al., 2019), focuses on searching for information, which is more relevant to web navigation than to the task of accessing a course. A reworded version removing "information" would better reflect the task at hand.

Item 12, originally measuring satisfaction with "This system has all the functions and potential corresponding to my expectations," faced issues based on expert feedback. The term "potential" was deemed inappropriate for an access procedure, and users had unclear expectations. Replacing "potential" with "effectiveness" still did not capture satisfaction adequately. A revised item, such as "The features available for accessing the course meet my expectations," would perhaps more accurately measure satisfaction, especially when combined with Item 11, which assesses interface pleasantness.

Rewording items 3 and 12 requires thorough re-evaluation to ensure they align with the usability construct. Content validation by subject matter experts and a new factor analysis are needed to verify the items' validity and ensure they contribute meaningfully to the overall measurement.

The factor loadings of items 3 and 12, while valuable, suggest an opportunity to enhance the robustness of the findings. Factor analysis benefits from a larger sample size, with recommendations typically ranging from 100 to 400 participants, depending on the number of variables and data characteristics (Guilford, 1956; Stevens, 2002). By expanding the sample size, future analyses could offer even more precise population estimations and more reliable inferences, further strengthening the validity of the results.

In summary, the study's findings suggest the following actions:

- Methodological Modification: Implement a 'usability testing with prompt' approach (Geisen & Bergstrom, 2017; Lazar et al., 2017; Shneiderman et al., 2017) to administer questions at critical navigation points, which will be evaluated in future surveys.
- Item Reformulation: Revise the problematic items to better align with the model's construct, improving accuracy and consistency. This revision will require expert validation and a new factor analysis with a larger student sample.

5. Conclusions

This study aimed to identify effective methods for assessing the usability of federated access to online courses in higher education mobility programs and to develop a validated instrument for academic institutions to evaluate the usability of their federated online offerings. The outcome is a tailored questionnaire created through an iterative process that included contextual analysis, instrument selection, adaptation, user feedback, re-adaptation, and final

validation. Participants completed a scenario-based task comprising two key steps: (a) authenticating via institutional credentials on a partner university's login page, and (b) locating and accessing T4EU courses on the partner university's Moodle platform.

The study also recommends adopting a "usability test with prompts" approach in future implementations to streamline question administration. Factor analysis revealed multiple dimensions consistent with the adapted NAU model, while also indicating the need for replication with larger samples. Some items did not fully capture key aspects of the latent variable, suggesting the need for targeted refinement in subsequent testing.

This work establishes a foundational instrument for evaluating the usability of interoperable Single Sign-On (SSO) LMS solutions. Its application to the T4EU federated Moodle—which continues to expand across partner institutions and courses—offers opportunities for further insights and comprehensive validation. The findings underscore the importance for T4EU member institutions to regularly use this instrument, as ongoing testing not only refines the evaluation process but also identifies specific areas for improvement within the Transform4Europe Moodle ecosystem, ultimately enhancing the user experience across the alliance.

Future research should explore variations in EFA scores across partner countries and academic disciplines, as well as factors influencing these differences, including users' ICT skills, fields of study, and prior experience with federated access interfaces. Additional usability dimensions—such as accessibility, device and browser compatibility, and support for users with disabilities—also warrant investigation. These efforts will strengthen the usability assessment framework and contribute to improving the overall experience for students engaging with federated access within a European University Alliance.

Acknowledgements

References

Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford.

Benmoussa, K., Laaziri, M., Khoulji, S., Kerkeb, M. L., & Yamami, A. E. (2019). AHP-based approach for evaluating ergonomic criteria. *Procedia Manufacturing, 32*, 856-863. https://doi.org/10.1016/j.promfg.2019.02.294

Berger, F., Galati, N., & Witteler, S. (2023). *Making interoperability work, challenges and solutions for an interoperable higher education system*. https://hochschulforumdigitalisierung.de/sites/defa ult/files/dateien/HFD_report_no.72_Making_inter operability_work.pdf

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246. https://doi.org/10.1037//0033-2909.107.2.238

Bentler, P. M., & Mooijaart, A. B. (1989). Choice of structural model via parsimony: A rationale based on precision. *Psychological Bulletin, 106*(2), 315–317.

Boudreau, M., Gefen, D., & Straub, D. (2001). Validation in IS research: A state-of-the-art assessment. *MIS Quarterly, 25*, 1–24.

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. Jordan, B. Thomas, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189-194). Taylor & Francis.

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research, 24*(4), 445–455.

Byne, B. M. (2010). *Structural equation modeling with Amos: Basic concepts, applications, and programming* (2nd ed.). Routledge-Taylor and Francis Group.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 213-218).

Chung, T. K., & Sahari, N. (2015). Utilitarian or experiential? An analysis of usability questionnaires. *International Journal of Computer Theory and Engineering, 7*(2), 167-171.

Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology, 55*(4), 584–594.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Erlbaum.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage Publications.

Diamantopoulos, A., & Siguaw, J. A. (2000). *Introducing LISREL*. Sage Publications.

Gaebel, M., Zhang, T., Stoeber, H., & Morrisroe, A. (2021). *Digitally enhanced learning and teaching in European higher education institutions*. European University Association absl.

Galende, B. A., Mayoral, S. U., García, F. M., & Lottmann, S. B. (2023). FLIP: A new approach for easing the use of federated learning. *Applied Sciences, 13*(6), 3446. https://doi.org/10.3390/app13063446

Geisen, E., & Bergstrom, J. R. (2017). *Usability testing for survey research*. Morgan Kaufmann.

Gilbert, G. E., & Prion, S. (2016). Making sense of methods and measurement: Lawshe's content validity index. *Clinical Simulation in Nursing, 12*(12), 530–531.

Gonzalez-Holland, E., Whitmer, D., Moralez, L., & Mouloua, M. (2017). Examination of the use of Nielsen's 10 usability heuristics & outlooks for the future. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 1472-1475). https://doi.org/10.1177/1541931213601853

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). W. B. Saunders.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103*(2), 265–275.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2010). *Multivariate data analysis* (7th ed.). Pearson.

Halim, E., Soeprapto Putri, N. K., Anisa, N., Arif, A. A., & Hebrard, M. (2021). Usability testing of vocabulary game prototype using the Nielsen's attributes of usability (NAU) method. In *Proceedings of the 2021 International Conference on Information Management and Technology* (pp. 590-594). https://doi.org/10.1109/ICIMTech53080.2021.9534970

Harman, H. H. (1976). *Modern factor analysis*. University of Chicago Press.

Hodrien, A., & Fernando, T. (2021). A review of post-study and post-task subjective questionnaires to guide assessment of system usability. *Journal of Usability Studies, 16*(3), 203-232.

Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. In R. Hoyle (Ed.), *Structural equation modeling: Issues, concepts, and applications* (pp. 76–99). Sage.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.

ISO. (2018). *Ergonomic requirements for office work with visual display terminals (VDTs)—Part 11: Guidance on usability* (ISO 9241-11).

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika, 35*, 401–415.Kass, N. L., & Tinsley, R. L. (1979). Factor analysis of psychological and educational data. *Journal of Educational Psychology, 71*(6), 844–855.

Kirakowski, J. (1995). Evaluating usability of the human-computer interface. In *Advances in Human-Computer Interaction: Human Comfort and Security* (pp. 21-32). Springer Berlin Heidelberg.

Kirakowski, J., & Cierlik, B. (1998). Measuring the usability of web sites. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 42, No. 4, pp. 424-428).

Latiar, H., Dwi, M., & Nining, S. (2024). Evaluation of repository usability test using Nielsen's attributes of usability (NAU) model in libraries. *Jurnal Ilmu Perpustakaan dan Informasi, 9*(1).

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In *Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society* (pp. 63-76). Springer Berlin Heidelberg.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563–575.

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann.

Lewis, B. R., Snyder, C. A., & Rainer, K. R. (1995). An empirical assessment of the Information Resources Management construct. *Journal of Management Information Systems, 12*, 199-223.

Lewis, J. R. (1992). Psychometric evaluation of the post-study system usability questionnaire: The PSSUQ. In *Proceedings of the Human Factors Society Annual Meeting* (Vol. 36, No. 16, pp. 1259-126

Lund, A. (2001). Measuring usability with the USE questionnaire. *Usability Interface, 8*, 3-6.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99.

Medsker, G. J., Williams, L. J., & Holahan, P. J. (1994). A review of current practices for evaluating causal models in organizational behavior and human resource management research. *Journal of Management, 20*, 439–464.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage Publications.

Munir, S., Rahmatullah, A., Saptono, H., & Wirani, Y. (2019). Usability evaluation using NAU method on web design technique for web portal development in STT Nurul Fikri. In *Proceedings of the 2019 Fourth International Conference on Informatics and Computing* (pp. 1-6). https://doi.org/10.1109/ICIC47613.2019.8985913

Newman, D. A. (2009). Missing data techniques and low response rates: The role of systematic nonresponse parameters. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity, and fable in the organizational and social sciences* (pp. 7–36). New York, NY: Routledge.

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods, 17*(4), 372–411.

Nielsen, J. (2012). How many test users in a usability study. Nielsen Norman Group. https://www.nngroup.com/articles/how-many-test-users/

Nielsen, J., & Kaufmann, M. (1993). *Usability engineering*. Morgan Kaufmann.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Osborne, J. L., & Costello, M. S. (2004). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation, 9*(1), 1–10.

Palmieri, P. A., Leyva-Moral, J. M., Camacho-Rodriguez, D. E., Granel-Gimenez, N., Ford, E. W., Mathieson, K. M., & Leafman, J. S. (2020). Hospital survey on patient safety culture (HSOPSC): A multi-method approach for target-language instrument translation, adaptation, and validation to improve the equivalence of meaning for cross-cultural research. *BMC Nursing, 19*, 1-13.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*, 489–497.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*, 459–467.

Ruoti, S., Roberts, B., & Seamons, K. (2015). Authentication melee: A usability analysis of seven web authentication systems. In *Proceedings of the 24th International Conference on World Wide Web* (pp. 916-926).

Sagar, K., & Saha, A. (2017). A systematic review of software usability studies. *International Journal of Information Technology*, 1-24. https://doi.org/10.1007/s41870-017-0048-1

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23–74.

Shneiderman, B., Plaisant, C., Cohen, M. S., Jacobs, S. M., & Elmqvist, N. (2017). *Designing the user interface: Strategies for effective human-computer interaction*. Pearson.

Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology, 5*, 1-25. https://doi.org/10.1146/annurev.clinpsy.032408.153639

Tabachnick, L., & Fidell, C. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson.

Taherdoost, H. (2016). Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. *International Journal of Academic Research in Management (IJARM), 5*.

Vlachogianni, P., & Tselios, N. (2023). Perceived usability evaluation of educational technology using the post-study system usability questionnaire (PSSUQ): A systematic review. *Sustainability, 15*(17), 12954. https://doi.org/10.3390/su151712954

Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development, 45*(3), 197–210.

Appendix: tables and figures

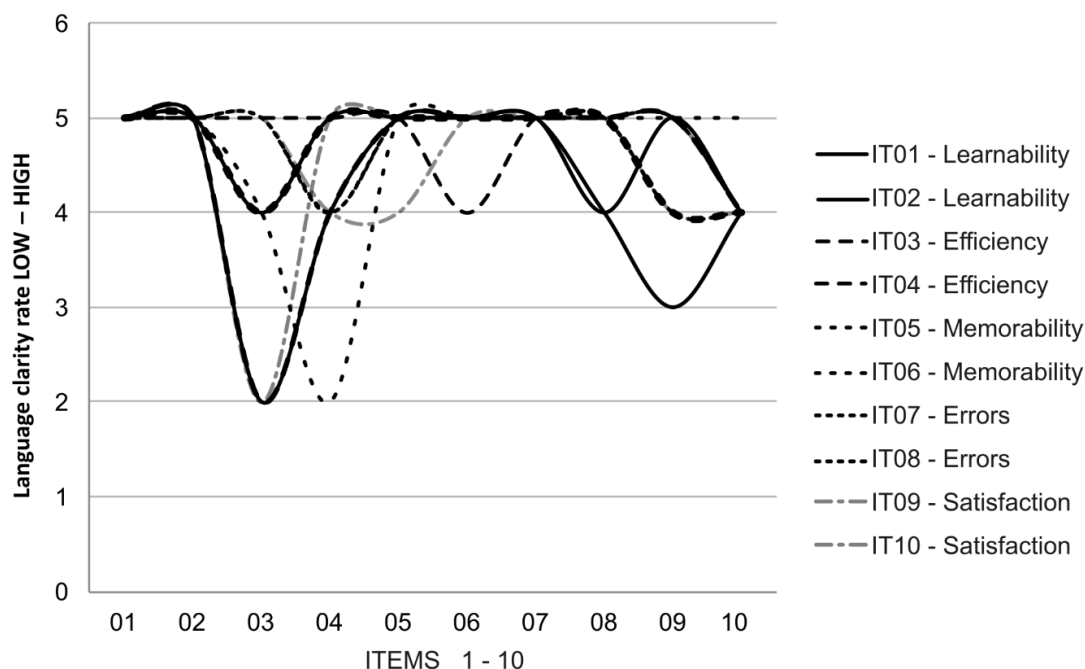**Figure 1** - Average rate of language clarity per item.

Table 1A - Translation and adaptations of the NAU-based questionnaire.
VERSION 1: after 1st translation and adaptation, administered to students in Phase 1.
VERSION 2: after 2nd translation and adaptation, administered to usability experts in Phase 2.

ATTRIBUTES	*ITEMS*	*VERSION 1*	*VERSION 2*	*N.*
Learning	This system is simple to use	Questa procedura di accesso è semplice da svolgere This access procedure is easy to perform.	Questa procedura di accesso al corso è facile da svolgere This procedure for accessing the course is easy to perform	1
	The information provided with this system is easy to understand	Le informazioni fornite durante la procedura sono facili da capire The information provided during the procedure is easy to understand	Le informazioni fornite durante la navigazione sono facili da capire The information provided during navigation is easy to understand	2
Efficiency	I easily found the information I am looking for	Ho trovato facilmente le informazioni che cercavo I easily found the information I was looking for	Ho trovato facilmente le informazioni che cercavo I easily found the information I was looking for	3
	I was able to quickly complete my task with this system	Svolgendo questa procedura sono riuscito/a ad accedere rapidamente al corso By carrying out this procedure I was able to quickly access the course	Sono riuscito/a ad accedere rapidamente al corso erogato dall'ateneo partner I was able to quickly access the course provided by the partner university	4

ATTRIBUTES	ITEMS	VERSION 1	VERSION 2	N.
Effectiveness			Le informazioni fornite durante la navigazione mi hanno permesso di accedere al corso The information provided during navigation allowed me to access the course	5
			Il layout delle schermate (icone, pulsanti, barre di navigazione, selezione della lingua e link) mi ha aiutato ad accedere al corso The layout of the screens (icons, buttons, navigation bars, language selection and links) helped me access the course	6
Memorability	The organisation of information in the system screens is clear	L'organizzazione delle informazioni nelle schermate è chiara The organisation of information in the screens is clear	L'organizzazione delle informazioni incontrate durante la navigazione è chiara The organisation of information encountered during navigation is clear	7
	The system is easy to remember	La procedura è facile da ricordare The procedure is easy to remember	La procedura di accesso al corso è facile da ricordare The course access procedure is easy to remember	8
Error	The error messages presented by this system tell me clearly how to solve problems	I messaggi di errore presentati nello svolgimento della procedura mi dicono chiaramente come risolvere i problemi The error messages presented in the procedure clearly tell me how to solve the problems	I messaggi di errore presentati durante la navigazione mi dicono chiaramente come risolvere i problemi The error messages presented during navigation clearly tell me how to solve the problems	9
	When I made a mistake using this system, it was easy and quick to correct it	Quando ho commesso un errore durante la procedura, è stato facile e veloce correggerlo When I made a mistake during the procedure, it was quick and easy to correct it	Quando ho commesso un errore durante la procedura di accesso al corso, è stato facile e veloce correggerlo When I made a mistake during the course access procedure, it was quick and easy to correct it	10
Satisfaction	The interface of this system is nice	Le interfacce per questa procedura sono piacevoli The interfaces for this procedure are nice	Le interfacce per questa procedura sono piacevoli The interfaces for this procedure are nice	11
	This system has all the functions and the potential corresponding to my expectations	Questa procedura di accesso ha tutte le funzioni e le potenzialità che corrispondono alle mie aspettative This access procedure has all the functions and potential to match my expectations	Questa procedura di accesso al corso ha l'efficacia e le funzioni che mi aspetterei This course access procedure has the effectiveness and functions I would expect	12

Table 2A - Content Validity Ratio.

	ITEMS	CVR per ITEM	CVI
Learnability	ITEM 1	1	
	ITEM 2	1	
Efficiency	ITEM 3	1	
	ITEM 4	1	
Effectiveness	ITEM 5	1	
	ITEM 6	1	0.91
Memorability	ITEM 7	1	
	ITEM 8	0.80	
Errors	ITEM 9	0.80	
	ITEM 10	0.80	
Satisfaction	ITEM 11	1	
	ITEM 12	0.60	

Table 3A - Anova.

	UNI	N	Mean	SD	SE
MEAN	USil	40	2.57	0.528	0.0834
	VMU	36	2.66	0.648	0.1081
	JMU	29	2.48	0.718	0.1333
	UA	35	2.44	0.705	0.1192

Table 4A - Skewness and kurtosis.

Items	N	Missing	Skewness	Kurtosis
1	140	0	-0.759	0.128
2	140	0	-0.693	0.148
3	140	0	-0.429	-0.350
4	140	0	-0.549	-0.401
5	140	0	-0.734	0.554
6	140	0	-0.576	-0.0501
7	140	0	-0.477	-0.091
8	140	0	-0.848	0.830
9	117	23	-0.460	0.0443
10	120	20	-0.450	0.0328
11	140	0	-0.518	0.275
12	140	0	-0.740	0.283

Table 5A - Exploratory Factor Analysis.

	1	2	3	4	5	6	Uniqueness
ITEM 1			0.578				0.37554
ITEM 2			0.833				0.17098
ITEM 3							0.37433
ITEM 4		0.975					0.00945
ITEM 5						0.526	0.39368
ITEM 6						0.600	0.29579
ITEM 7					0.772		0.20622
ITEM 8					0.400		0.60186
ITEM 9	0.967						0.10034
ITEM 10	0.615						0.40322
ITEM 11				0.985			0.00500
ITEM 12							0.32645

Table 6A - Factor Loadings.

Factor	SS Loadings	% of Variance	Cumulative %
1	1.702	14.18	14.2
2	1.535	12.79	27.0
3	1.639	13.66	40.6
4	1.406	11.72	52.4
5	1.523	12.69	65.0
6	0.932	7.77	72.8

Table 7A - Proposed rewording for items 3 and 12.

Attribute		ITEM	Reformulation
Efficiency	3	Ho trovato facilmente le informazioni che cercavo I easily found the information I was looking for	Ho trovato facilmente quello che cercavo I easily found what I was looking for
Satisfaction	12	Questa procedura di accesso al corso ha l'efficacia e le funzioni che mi aspetterei This course access procedure has the effectiveness and functions I would expect	Le funzionalità disponibili per l'accesso al corso corrispondono alle mie aspettative The features available for accessing the course meet my expectations.